

NEWSPAPER READING IN THE SECONDARY MODERN SCHOOL

(AN ENQUIRY INTO THE READING INTERESTS OF PUPILS IN THIRD-YEAR CLASSES IN SECONDARY MODERN SCHOOLS.)

By DONALD G. BURNS

Lecturer in Education, University of Leeds.

I.—*Introductory.* II.—*Newspapers read by these pupils.* III.—*Classification of newspaper content.* IV.—*What portions of the newspaper do boys and girls read?* V.—*Discussion.* VI.—*Summary of results.*

THE general development of children's reading interests has been traced in a number of articles published in this *Journal* and elsewhere, but little information is available on what children actually read in the comics, adventure papers and other reading matter which they mention in their replies to questionnaires. The enquiry carried out by Jenkinson¹ in 1938, provides information on the reading interests of children in grammar schools and the A streams of senior schools but, with the possible exception of Chapters VII and XXI, contains no direct information on the material actually read. Moreover, these findings are based on a selected portion of the school population and significant changes in children's reading interests may have taken place since the date of that enquiry. A more recent enquiry directed by Miss Stewart² amongst senior school children in Ilford, in 1946, provides information on the popularity of many periodicals and newspapers among children in the age-groups 12-15 in relation to their other leisure activities and the subject has been further developed by Williams,³ but it was not the purpose of these investigations to show which portions of periodicals are most widely read by children. Information on the portions of the newspaper read by different groups of adolescents is given in a detailed analysis by Wall⁴ though changes may be expected to have occurred in the popularity of some items since 1945. "In most groups," says Wall, "On the average, Current News of the War stands highest both in numbers of those reading it and in interest score." (*op cit.* p. 36) and this item has virtually disappeared from the daily paper to-day.

The present article deals only with the newspaper reading of boys and girls in the secondary modern school and is based on the answers to a questionnaire in which pupils were invited to state which newspapers they read.⁵ The enquiry was undertaken in May, 1953, to obtain detailed information about the reading interests of pupils in the third year classes of secondary modern schools, the latest stage for which results can be obtained for a complete academic year. Nine schools co-operated in the enquiry and these were chosen as far as possible to avoid weighting in favour of one section of the population. They included three boys', three girls' and three mixed schools and are situated in lower middle

¹ JENKINSON, A. J.: *What Do Boys and Girls Read?* (Methuen, 1940.)

² STEWART, M.: *The Leisure Activities of School Children.* (W.E.A., 1948.)

³ WILLIAMS, A. R.: "The Magazine Reading of Secondary School Children." (This *Journal*, XXI (iii), p. 186.)

⁴ WALL, W. D.: "The Reading of Newspapers among Adolescents and Adults." (This *Journal*, XVIII (i), p. 26.)

⁵ A part only of the questionnaire is dealt with here: other questions concerned the reading of periodicals and books.



2 Newspaper Reading in the Secondary Modern School

class and working class districts with the exception of one mixed school—a selective modern school situated in a residential area. All pupils in third year classes were asked to complete a questionnaire and replies were received from 415 boys and 383 girls.¹ Interviews were then arranged with a number of pupils chosen at random from each class in four of the schools and pupils were asked to look at the pages of the newspapers mentioned in their replies to the questionnaire and name the sections which they read regularly. A record was made of interviews with forty-five boys and thirty-two girls.

TABLE I
NEWSPAPER READING OF A GROUP OF CHILDREN IN SECONDARY MODERN SCHOOLS,
AGED 13-14+

Title	Boys (N=415)		Girls (N=383)	
		Per cent.		Per cent.
Local Daily :				
<i>Yorkshire Evening Post</i>	252	60	228	59
<i>Yorkshire Evening News</i>	128	30	95	24
<i>Yorkshire Post</i>	52	12	42	10
National Daily :				
<i>Daily Mirror</i>	82	19	75	19
<i>Daily Express</i>	58	14	42	10
<i>Daily Mail</i>	35	8	17	4
<i>Daily Herald</i>	33	8	47	12
<i>News Chronicle</i>	14	3	15	4
<i>Daily Sketch</i>	—	—	6	1
Weekly :				
<i>News of the World</i>	106	25	93	24
<i>The People</i>	70	16	59	15
<i>Sunday Pictorial</i>	35	8	70	18
<i>Titbits</i>	18	4	14	3
<i>Empire News</i>	18	4	20	5
<i>Sunday Dispatch</i>	16	3	13	3
<i>Overseas Mail</i>	9	2	16	4
<i>Sunday Chronicle</i>	7	1	—	—
<i>Sunday Express</i>	5	1	8	2

Newspapers mentioned less than five times are excluded.

II.—NEWSPAPERS READ BY THESE PUPILS.

Table I shows the daily and weekly papers most widely read by children in these nine secondary modern schools. The papers which are more widely read than any others are the *Yorkshire Evening Post* and the *Yorkshire Evening News*, neither of which is mentioned in the tables of the Hulton Readership Survey.² The tendency for boys and girls in the provinces to read a higher proportion of local daily papers than those living nearer London has been noted in other enquiries (cf. Wall *op. cit.*), and as these are both evening papers, it may be assumed that they occupy a complementary position with the national dailies. These—*The Daily Mirror*, *The Daily Express*, *The Daily Mail* and *The Daily Herald*—appear in the same order here as in the Hulton Readership Survey (Table I).² This was to be expected since it is doubtful whether any

¹ The age distribution was as follows :

	13+	14+
Boys	200	215
Girls	162	221

² *The Hulton Readership Survey* (1953). (Hulton Press, Ltd.)

school children buy newspapers regularly and they will choose their reading from the papers bought by adults.

Of the weekly papers read by these children, the most popular—*The News of the World*, *The People* and *The Sunday Pictorial* appear in the same order here as in the Hulton Survey of adult reading (Table I). It was noted that less than five children mention *The Radio Times* which is given as the most widely read of weekly papers in the Hulton Child Readership Survey¹ and second only to *The News of the World* in the survey of adult reading (*op. cit.* Table I).

Most pupils named several daily and weekly papers in reply to the question "What daily and weekly newspapers do you read regularly?" and each boy and girl named an average of 1.5 daily papers. It may be doubted whether this number is read by each pupil every day, but it is of interest to note that approximately the same averages were obtained by Wall² for a smaller group of non-grammar school pupils in 1948 (1.6 for boys and 1.5 for girls) (*op. cit.* p. 31, Table 1A).

III.—CLASSIFICATION OF NEWSPAPER CONTENT.

None of these newspapers martials the news from home and abroad, the editorial comment, the advertisements commercial and private and sporting events in the orderly manner of *The Times*. In some papers (e.g., *The News of the World*), the editor's policy is plainly to offer items from most of these topics on every page. Nor did the analysis of newspaper content adopted in other enquiries seem likely to offer information in sufficient detail about the portions of the newspaper which children read. Jenkinson's threefold analysis into News, Sport and Children's Column (*op. cit.* p. 81) is, as he suggested, "Perhaps too comprehensive to be of much value in itself." "The news includes everything except the sport and the children's column so that this does not reveal at all precisely what (they) read" (p. 234).

Wall's analysis of newspaper content into thirteen sections (*op. cit.* p. 35), although it reveals very much more about the newspaper reading of adolescents, did not seem likely to provide sufficient information about the reading of this somewhat younger group. Few entries could have been recorded, for example, under "Current News of the War" and "Military Correspondent," whereas a large number could be made under "Gossip" and "Domestic News." Here it seemed that a provisional distinction at least should be made between (a) news which is topical but not sensational (Korea, the Coronation and school sports); (b) news of the most sensational kind (murders, executions and assaults); and (c) news commentary.

Some distinction could also be made between the reading of commercial advertisements (some of which occupy half a page or more), the small shopping advertisements and classified advertisements of the type "Houses for Sale" and "Lost and Found."

Again the rubric "Sports News" seemed unduly comprehensive since it may cover two or even three pages of the newspapers read by these children and it seemed worth while to draw a distinction here between sports results and news about sports.

¹ *The Hulton Child Readership Survey* (1950). (Hulton Press, Ltd.)

² STEWART states in her survey of the *Leisure Activities of School Children* (W.E.A.): "Apart from the *Daily Mirror*, no paper was mentioned by an appreciable number of boys and girls. Only 34 per cent. of the boys and 31.7 per cent. of the girls included a newspaper among the papers and periodicals which they liked to read" (p. 16).

Boys and girls in grammar schools were reported to read an average of 1.8 newspapers at age 13+: STEWART, "Leisure Activities of Grammar School Children." (*This Journal*, XX (i), p. 11.)

4 Newspaper Reading in the Secondary Modern School

The following is the classification which was finally adopted :

- (i) Front page.
- (ii) Other news (including sensational items and references to photographs).
- (iii) Gossip and comment (including editorial) ; (correspondence).
- (iv) Advertisements (commercial) ; (shopping) ; (classified).
- (v) Cartoons (political) ; (comic) ; (strip) ; (sports).
- (vi) Sports (cricket only) ; (all sports pages) ; (articles on other pages).
- (vii) Miscellaneous (pets and animals) ; (crosswords) ; (T.V. and radio) ; (astrology) ; (weather)¹.

Information on pupils' reading interests under each of these headings was obtained by interviewing 45 boys and 32 girls chosen at random from 14 classes in four of the schools where the questionnaire had been answered. Each pupil was asked to show by reference to four daily papers and two weeklies which portions of the newspaper he read regularly. Boys were questioned on *The Yorkshire Evening Post*, *The Yorkshire Evening News*, *The Daily Mirror*, *The Daily Express*, *The News of the World*, and *The People* : in accordance with the preferences shown in answers to the questionnaire (Table I), *The Daily Herald* and *The Sunday Pictorial* replaced *The Daily Express* and *The People* in interviews with girls.

The following is a summary of the statements recorded for each newspaper :

Title	Boys	Girls	Total
<i>Yorkshire Evening Post</i>	25	17	42
<i>Yorkshire Evening News</i>	9	14	23
<i>Daily Express</i>	12	—	12
<i>Daily Mirror</i>	11	16	27
<i>Daily Herald</i>	—	10	10
<i>News of the World</i>	22	21	43
<i>The People</i>	14	—	14
<i>Sunday Pictorial</i>	—	19	19
	93	97	190

IV.—WHAT PORTIONS OF THE NEWSPAPER DO BOYS AND GIRLS READ ?

These Tables show that 43 per cent. of the items mentioned by boys and 38 per cent. of those mentioned by girls are news items of one kind or another. All of those questioned for example, claim to read the front page of *The Yorkshire Evening News* and the majority claimed to read all of it. Moreover, it was apparent in most cases that mention of a news item meant more than glancing through it—it meant that some of the article had been read. Only 51 of the 479 items mentioned by boys carried an illustration and 55 of the 513 mentioned by girls (in each case, just over 10 per cent.) and in view of the anxiety frequently expressed about pupils' ability to read when they leave school, it seems of some interest that a majority of the items mentioned under this heading must be read to be understood.

Further, the number of references to sensational items was smaller than expected. It is possible, however, that more are interested in these items for all those recorded were the pupils' spontaneous comments. Some boys explicitly disclaimed any interest in reports of murders or assaults with violence and those who did mention their interest were quite explicit. "On page two, I read murders and anything like that" (B). "On page three, I read the advertisements and the crime, if there is any : there generally is, and then I read the whole page" (F).

¹ An alternative classification is suggested by Lazar (quoted WITTY in *Reading in Modern Education*, p. 41). The contents of American newspapers are divided into comics ; sport ; news ; crimes ; stories ; crossword puzzles ; front page ; all ; editorial ; rotogravure ; magazine ; political. This did not seem as likely to discriminate between the items in British newspapers as the classification suggested here.

After news, advertisements constitute one of the principal features of newspaper reading for these children, and the number of references under this heading was surprisingly high (18 per cent. of all items mentioned by boys and 20 per cent. of those mentioned by girls). Of the large commercial advertisements, the most popular with boys and girls appeared to be the illustrated advertisement for forthcoming films in *The News of the World* (12 per cent. of mentions from boys refer to this; 7 per cent. from girls); but most of the commercial advertisements in the newspapers examined seemed to have a potential interest for both sexes and none were passed over without mention.

The small shopping advertisements which appear in Sunday papers also seem to make a strong appeal and approximately one-fifth of boys' and girls' references to advertisements in these papers mention shopping advertisements (18 per cent. for boys; 20 per cent. for girls).

The mentions under classified advertisements are, perhaps, more significant than they appear from Tables II and III for they appear only in local papers

TABLE II
BOYS' READING OF NEWSPAPERS: SECTIONS RANKED ACCORDING TO FREQUENCY OF MENTION.¹

Section	Y.E.P.	Y.E.N.	Daily Mirror	Daily Express	People	News of World	Sub-Total	TOTAL
I.—FRONT PAGE	25	9	11	12	14	22	93	
II.—OTHER NEWS	96	29	44	60	72	85	386	479 (43%)
(Sensational items) . .	16	1	0	2	9	2	30 (6%)	
(References to photos.)	29	0	5	8	9	0	51 (10%)	
III.—GOSSIP AND COMMENT	36	15	0	5	6	16	78 (96%)	
Correspondence	0	—	3	—	—	0	3 (4%)	81 (7%)
IV.—ADVERTISEMENTS								
Commercial	21	4	10	19	14	76	144 (71%)	
Shopping	—	—	—	—	14	6	20 (9%)	
Classified	21	16	—	—	—	—	37 (18%)	201 (18%)
V.—CARTOONS :				8	—	—	8 (4%)	
Political	—	—	—	11	0	17	91 (49%)	
Strip	—	0	63	0	0	0	16 (9%)	
Sports	—	8	8	0	0	0	68 (37%)	183 (16%)
Comic	12	4	8	15	17	12		
VI.—SPORT :								
Cricket only	8	3	7	3	12	0	33 (34%)	
All sports page	11	5	0	4	0	0	20 (20%)	
Articles on other pages	—	—	6	0	9	29	44 (45%)	97 9(%)
VII.—MISCELLANEOUS :								
Pets and Animals	5	0	0	0	0	6	11 (15%)	
Crosswords	—	1	3	0	5	14	23 (32%)	
T.V. and Radio	17	2	5	0	0	0	24 (33%)	
Astrology	—	—	—	—	4	6	10 (14%)	
Weather	0	0	0	4	0	0	4 (5%)	72 (6%)

A dash indicates that the item did not appear in the papers examined.

¹ For convenience in making comparisons between the sections, the number of references in each section has been given as a percentage of the total. The total for sub-sections has been given also as a percentage of all references recorded within the section though some of the samples are very small at this level.

6 *Newspaper Reading in the Secondary Modern School*

and may be read several times a week though listed only once. The serious purpose behind a good deal of the reading of these portions of the newspaper was sufficiently indicated by the sections which boys and girls mentioned. Boys referred to "looking down the jobs," "Workpeople," "Sales by Private Contract," and girls to "Typists," "Rooms to Let," "Jobs of all Kinds," "Clerks" and "Workpeople." But their interests were by no means limited to these topics and frequent reference was made to "Births and Deaths" ("to see if there is anyone I know"), "Lost and Found," "Seaside Resorts," "Cars," and "Motor Caravans."

Cartoons of all kinds form the next most popular item. Only one of the six papers examined with these boys publishes a political cartoon and three of the eight reading it doubt whether they understood it. But all seemed to read and enjoy the small cartoon with a single drawing embedded in a column of print, and many showed by the readiness with which they named it or found it in the paper, that they were in the habit of reading this item.

The strip cartoon may assume undue prominence in these Tables owing to its appearance as a major feature in *The Daily Mirror*, where a whole page is devoted to six separate strips: but this feature is undoubtedly very popular and 67 per cent. of the girls' references to cartoons mention this. Not all the

TABLE III
GIRLS' READING OF NEWSPAPERS: SECTIONS RANKED ACCORDING TO FREQUENCY OF MENTION.

Section	Y.E.P.	Y.E.N.	Daily Mirror	Daily Herald	News of World	Sunday Pictorial	Sub-Total	TOTAL
I.—FRONT PAGE	17	14	16	9	21	19	96	
II.—OTHER NEWS: ..	91	33	106	22	72	93	417	
(Sensational Items) ..	15	22	1	5	0	11	54 (10%)	
(Photographs)	32	1	12	3	2	5	55 (10%)	513 (38%)
III.—GOSSIP AND COMMENT	33	8	0	0	19	68	128 (91%)	
Correspondence	4	0	7	0	0	2	13 (9%)	141 (10%)
IV.—ADVERTISEMENTS: ..								
Commercial	37	14	39	3	73	30	196 (72%)	
Shopping	—	—	—	7	10	16	33 (12%)	
Classified	18	24	—	—	—	—	42 (15%)	271 (20%)
V.—CARTOONS: ..								
Political	—	—	—	—	—	—	0	
Strip	—	—	98	10	19	32	159 (67%)	
Sports	—	5	3	10	9	—	27 (11%)	
Comic	10	10	10	8	—	12	50 (21%)	236 (17%)
VI.—SPORTS: ..								
Cricket only	6	2	0	0	6	4	18 (25%)	
All sports page	3	1	10	4	5	—	23 (32%)	
Articles on other pages ..	—	0	4	3	19	4	30 (42%)	71 (5%)
VII.—MISCELLANEOUS: ..								
Pets and animals	4	0	0	0	21	0	35 (28%)	
Crosswords	—	10	6	3	12	—	31 (26%)	
T.V. and Radio	18	9	4	2	0	4	37 (30%)	
Astrology	—	—	—	—	11	2	13 (10%)	
Weather	0	0	0	6	0	0	6 (5%)	122 (9%)

A dash indicates that the item did not appear in the newspaper examined.

strips, however, are read by all the girls (85 per cent. of recorded readings) or boys (75 per cent. of recorded readings) and for both sexes, the most childish strip (Pip, Squeak and Wilfred) was the least popular. In general, girls seem to show a greater preference for the strip cartoon than boys, a tendency already noted by Wall (*op. cit.* p. 94).

Sports rank fourth for boys (9 per cent. of all mentions)¹ and seventh for girls (5 per cent. of all mentions) and nearly half the mentions refer to articles about sport in the two Sunday papers. The remainder refer explicitly either to the results and comment on cricket or to the whole of the 'sports' page. It can hardly be doubted that the general interest in cricket will have been heightened by the presence of the Australian team in England in June (when these replies were recorded), though it was impossible to estimate what effect this may have had. Other differences between the sexes are apparent in the balance of reading interest between the sports sections. Girls seem to show a more general interest in sports than boys and 32 per cent. of the girls' replies refer to news about sports against 20 per cent. of the boys' answers. The boys' answers suggest a more specific interest in these pages and 34 per cent. refer to cricket only (girls 25 per cent.). Nearly half the reading about sports is found by both sexes in articles in other pages (45 per cent. for boys and 42 per cent. for girls).

Girls seem to show a far greater interest in the 'gossip' and 'comment about news' columns and these they read more frequently than the sports pages. One of the Sunday papers most widely read by girls (*The Sunday Pictorial*) edits a large portion of the news in the form of comment and some of this apparent preference for 'comment' may arise from the choice of paper. Whatever the cause, the girls in this group evidently read far less than the boys about sport and a good deal more comment about the news. Gossip, comment and readers' letters rank fourth among the items read by girls (10 per cent. of all replies) and fifth for boys (7 per cent. of replies). Most of the reading under this heading consists of comment on the news and very few references are made to the correspondence columns and readers' letters. Not one pupil mentioned the editorial in any of the papers read and the small differences which were recorded between boys' and girls' reading of these portions of the newspaper suggests that there is little differentiation between the sexes at this stage.

Other items which pupils recorded as reading but which could not be suitably included in one of the major groups are crosswords (which many pupils said they tried if not too difficult), articles dealing with animals or pets (only explicit references were recorded and it is probable that many more read them) and occasional references to the weather, lighting-up time, radio programmes and astrology.

V.—DISCUSSION.

Some boys found it difficult to proceed from page to page in a newspaper and seemed inhibited if asked to detail what they read. Thus, S. said: "I read the whole of the first page and interesting parts in other pages. I read Jane and I read the middle pages. I look at Buck Ryan and do the crossword puzzle. I always look at Jack Dunkerly and I read the weather and lighting-up time." The statements of these boys—three only—were taken down and interpreted later.

In general, boys seemed selective in their newspaper reading. None claimed to read any paper from cover to cover and the majority turned without hesitation to the features they usually read, which seemed to suggest they had fairly

¹ WALL writes (*op. cit.* p. 97): "Sports News reaches a high level of interest among boys and men at all ages and is of comparatively minor importance to girls and women." It is of interest to note, however, that in that enquiry the boys in the technical and elementary group (aged 13.0 to 16.11) rank Sports News lower than any other male group (Table II, p. 35) and that the item appears fourth among the thirteen sections tabulated.

well-defined habits of reading. This seemed to be true for all the pupils interviewed whatever their level of academic attainment: thus, H. said, referring to *The Yorkshire Evening Post*, "I read the headlines on the front page, 'It happened to me' (comment on the news), 'Muffin the Mule' (a cartoon) and 'Angler's Scrapbook' every Friday. That's all."

The close relation between these reading habits and pupils' present interests may be gathered from the following summaries of two statements made in the course of an interview:

- (a) H. reads in *The Yorkshire Evening News* all the front page, news of 'general interest' and advertisements on page two, any news from collieries on page three, the 'News gossip' and cartoon on page four, commercial advertisements on page five and he goes through the classified advertisements on page six for "Lost and Found," "Caravans and Motor Cars for Sale," and "Workpeople."
- (b) V. reads, in *The Yorkshire Evening Post*, all the front page, on page two "Auctions," and "Lost and Found," nothing on page three, pets and birds on page four, advertisements about sales at C. and A.'s and Lewis's on page five, motor cars on page six, the Korean war on page seven, all of page eight, nothing on pages nine and ten, the illustrations on page eleven, advertisements for motor-cycles, cars and cycles on page twelve, holiday resorts on page thirteen, and the sports on page fourteen.

There was evidence from the girls also that many portions of the newspaper were 'skipped' and those items preferred were mentioned so consistently that it was often possible to anticipate a pupil's choice as she glanced through the paper. Photographs and references to the Queen, for example, almost always attracted attention. But here, too, the 'skipping' appeared to be methodical for girls pointed without hesitation to the items which they read, however inconspicuous they might be. Moreover, the replies of individual pupils suggested a wide range of interests. Some impression of this catholic taste may be gathered from the following summary of two interviews with girls:

- (a) C. reads, in *The Yorkshire Evening News*, the headlines on the front page and any news item about the Queen: advertisements, news of accidents and murder trials (if any) on page two; the headlines only on page three; the cartoon, radio programmes and 'Favourite Pets' on page four; the captions to photographs and a news item about the "Schoolgirl Queen" on page five; advertisements for holiday resorts, office jobs, animals for sale, and cinemas on page six; Births and Deaths on page seven; and the sports cartoon on page eight.
- (b) M. reads, in *The Yorkshire Evening Post*, about the ascent of Everest (page one); advertisements (page three); a news item about dogs (page four); commercial advertisements (page five); news about atom bombs (page six); news about Everest (page seven); the cartoon and cinema advertisements (page eight); a news item—'the Jealous Husband'—(page nine); news of Children's Day (page ten); the captions under illustrations (page eleven); advertisements for 'Rooms to let' (page twelve), and the Rugby football results (page fourteen).

VI.—SUMMARY OF RESULTS.

1.—The results of this enquiry suggest that the habit of newspaper reading may be more widespread among children in secondary modern schools than had been supposed and that for some, it is a well-established habit at the age of 13 plus. This seems to be equally true of boys and girls for both sexes named an average of 1.5 daily papers. Girls named slightly more Sunday newspapers than boys (.9 against .8) though the number of cases is too small to warrant the suggestion that they read them more regularly than boys.

Two local papers are the most widely read of the daily papers and after these, the four national papers which are most widely read throughout the country as a whole—*The Daily Mirror*, *The Daily Express*, *The Daily Mail*, and *The Daily Herald*. Mentions of Sunday papers show that *The News of the World*, *The People* and *The Sunday Pictorial* rank in the same order of popularity with these children as they do amongst adult readers.

2.—The general trend of pupils' interests in these newspapers is similar for both sexes except for sports. For boys, news of all kinds ranks first, advertisements second, cartoons third, sports fourth, gossip fifth, and "Miscellaneous" last; girls show the same preferences in their reading, with the exception of sports items, which are the least popular.

3.—Reading of news items appears to occupy a larger place with boys (43 per cent.) than it does with girls (38 per cent.). Virtually every pupil claimed to read news items on the front page but boys read appreciably more news than girls on other pages of the paper. On the other hand, 10 per cent. of the girls' replies on news referred to "sensational items" as against 6 per cent. of replies from boys.

4.—Advertisements rank second to news in order of popularity, being mentioned in one-fifth of the girls' replies and just under one-fifth of the replies from boys. Large commercial advertisements occupy the greater part of the advertising space in these newspapers and are mentioned as frequently by boys as girls (71 per cent. and 72 per cent. of replies on advertisements). The small shopping advertisements which appear principally in the Sunday papers are mentioned more frequently by girls than boys (12 per cent. against 9 per cent. for boys) and the classified advertisements more frequently by boys (18 per cent. against 15 per cent. for girls).

5.—Cartoons rank third in popularity and are mentioned almost as often by boys as girls; but under this general heading there were marked differences between the sexes. Girls showed a greater preference for strip cartoons than boys (67 per cent. of girls' answers; 49 per cent. of boys.), and boys seemed to show a preference for the comic or single picture cartoon (37 per cent. against 21 per cent. of answers from girls). Sports cartoons on the other hand seemed to be almost as popular with girls as with the boys.

6.—Sports items are mentioned less frequently than was expected (cf. Wall., *op. cit.* p. 97) and rank fourth in preference for boys (9 per cent. of all answers) and last with girls (5 per cent. of all answers).

7.—Gossip, comment and readers' letters rank fourth amongst items read by girls (10 per cent. of all replies) and fifth for boys (7 per cent. of all replies). The majority of the items referred to here consist of comment on the news; few refer to the correspondence columns (girls, 9 per cent.; boys, 4 per cent.) and none mentioned the editorial.

8.—The variety of small items grouped under miscellaneous is a further indication of the wide range of interests which children appear to satisfy in reading the newspapers. TV and radio programmes are mentioned more frequently than any other item in this group and other items, such as astrology, which do not appear in all papers, may be more popular than is shown by these results. Items about pets and animals are read by boys and girls though the girls appear to read these more regularly than the boys (28 per cent. against 15 per cent.); and boys show a greater interest in crossword puzzles.

ACKNOWLEDGEMENTS.

The writer wishes to thank Professor Valentine for a number of helpful suggestions and the Leeds Education Authority for kindly allowing the enquiry to be made in the Authority's schools and he is particularly grateful to the head teachers in the schools concerned for their willing co-operation.

THE PERMANENT CONTRIBUTIONS OF McDUGALL TO PSYCHOLOGY*

By CYRIL BURT

I.—*Early Education*. II.—*Experimental Research*. III.—*Cognitive Psychology : Child Study and Mental Tests*. IV.—*Conative Psychology : Instincts, Sentiments, and Social Life*. V.—*Body and Mind*.

I.—EARLY EDUCATION.

WILLIAM McDougall was born at Chadderton, a small town near Manchester, in 1871. He was the son of a wealthy manufacturing chemist, who sent him first to a private school nearby and then to a *Realgymnasium* at Weimar. At the early age of fifteen he entered the University of Manchester, where he obtained First Class Honours in Science. At nineteen he went to St. John's College, Cambridge, and there gained a First Class in both parts of the Tripos. He had already resolved "to make psychology his life's vocation," and was fully convinced that the proper approach no longer lay through philosophy—the path pursued by his academic predecessors, but ought to start from a first-hand knowledge of the anatomy and physiology of the brain. Accordingly, since courses in these subjects were confined to medical students, he registered in the faculty of medicine. As soon as he was qualified, he was elected a Fellow of his college, and in the following year was invited to join the Cambridge Anthropological Expedition to the Torres Straits. Here he helped Rivers and Myers in applying psychological tests to natives, and then spent six months studying social life among the head-hunters of Borneo (3). On returning to Europe, he stayed for a year at Müller's laboratory in Göttingen "to secure a more thorough grounding in psychophysical techniques."

These early experiences did much to determine his later attitude towards the problems of mind. As a schoolboy he acquired a facility in German—the language in which most of the experimental research in psychology was then published. As a medical student he learnt to attack psychological issues with an up-to-date knowledge of neurology, and acquired a keen interest in psychiatry—at that time a comparatively new branch of medical science. St. John's College† was already regarded as "a home for budding psychologists". G. F. Stout had just been re-elected to his Fellowship; W. H. R. Rivers was Tutor and Praelector; Sherrington a frequent guest. Finally, McDougall's own adventures among the Malay aborigines gave him an insight into human nature, and "a first-hand appreciation of the profound influence exercised by local patterns of culture on the thoughts and behaviour of the individual."

*In preparing this survey of McDougall's work, I have occasionally drawn on the obituary notice compiled by Dr. C. S. Myers and myself for the *Psychological Review* and on the "Introduction" I was asked to write by the Oxford University Press for the last edition of McDougall's book on *Psychology, the Study of Behaviour*. Here and there I have made use of extracts from his letters which seemed to throw light on the aims he had in view, particularly at the early stages of his life. For further biographical details, the reader may refer to "William McDougall: An Appreciation," this JOURNAL, IX, 1939, pp. 1-7.

†Some years later, when I went into residence there on accepting a post in the Psychological Laboratory at Cambridge, I heard, largely from the lips of Udney Yule, then Reader in Statistics, a string of amusing anecdotes about their lively debates on "body and mind" and the horror occasioned in the Senior Common Room when Rivers or McDougall would suddenly fetch "a defunct human brain or the eye of a dead ox to exemplify their theories to Stout."

In 1897 Professor Sully decided to establish a laboratory for experimental psychology at University College, London—the first of its kind in the British Isles; and a few years later McDougall was appointed Director in succession to Rivers. In 1903 Oxford invited him to follow Stout as Wilde Reader in Mental Philosophy; and for the next four years he held the two posts concurrently. In 1907 he resigned his London post, and so found leisure to write the books by which he is chiefly remembered.

II.—EXPERIMENTAL RESEARCH.

It was McDougall's firm conviction that, if British psychology was to be converted from a branch of philosophy into a branch of natural science, the first step was to persuade his fellow countrymen that it was a fit and urgent subject for experimental research. In this he encountered an active opposition that today must seem incredible. Accordingly he resolved to convert his own attic into a laboratory, and then begin with the lowest levels and work up to the highest.

His earliest articles, published in *Mind*, recorded a remarkable series of experiments on the nature of colour vision (4). Rivers had just published, in Vol. II of Schäfer's *Textbook of Physiology*, an erudite chapter on vision, reviewing the work done up to date, but leaving any decision about colour theories still an open question. At that time the prevailing theory was that of the German physiologist, Hering, who assumed the existence of three antithetic pairs of primary colours—white and black, red and green, blue and yellow. McDougall's observations on colour-blindness had already induced him to prefer a much older hypothesis, put forward a century before by the English physicist, Thomas Young—a versatile investigator to whom we owe not only the term "energy" but also the early decipherment of Egyptian hieroglyphics. Young's theory recognized only three primary colours—red, green and blue, and, as both Clark Maxwell and Helmholtz had shown, offered the most natural interpretation of the facts of colour mixture. To decide between the two rival theories McDougall devised a number of ingenious experiments; and, to account for certain phenomena unknown to Young, he supplemented the older version by suggesting that colour-vision in man was to be regarded as the outcome of a complex evolutionary process. In all its essential features McDougall's own interpretation is the one that prevails today.

Two years later he turned to a more ambitious study on the "nature of the attention process" (5). Here it is Stout's influence that is most conspicuous. Stout had sought to substitute for the traditional laws of association a broader principle of mental organization (or "noetic synthesis" as he termed it) which enabled him to offer a truer interpretation of such cognitive processes as "relative suggestion" (i.e., the "eduction of correlates"), "schematic apprehension," and the peculiar phenomena to which Ehrenfels had just given the name "Gestalt"—all of which he explained as forms of "apperception" (cf. 2, Vol. I, pp. 65f.). To McDougall "noetic synthesis" appeared to be the conscious aspect of what Sherrington had described as "the integrative action of the nervous system"; and his primary aim was to suggest an appropriate neurological mechanism for Stout's theories and to support the explanation by a series of laboratory experiments. The results led him to elaborate the view that the essential basis of intellectual activities consists in a hierarchy* of

*It is instructive to note that writers who have recently tried to design artificial mechanisms that will reproduce some of the distinctive features of human behaviour have also been led to propose a "hierarchy of organizers": cf., for example, D. M. MacKay, "Supralogical Behaviour in Automata," *Proc. Internat. Congr. Psychol.*, 1954, pp. 12 and 18.

neural circuits (or "arcs") arranged in four or five main levels—reflex, sensori-motor, perceptual, associative, and relational. As a consequence partly of an intricate architecture inherited at birth and partly of individual learning, these circuits were assumed to be organized into distinctive patterns or systems, which he christened (rather clumsily) "psychophysiological dispositions."* Liberated by external stimuli, a special form of "psychophysical energy" was then supposed to flow through these pre-existing circuits, guided by the relative resistances at the nerve-junctions. Attention was thus aroused; and the individual thereupon "apperceived" the stimuli in terms of his previous experience. The higher cell-junctions themselves were depicted as the probable seat of consciousness;† and the whole description was illustrated and confirmed by novel experiments on the apperception of ambiguous diagrams and pictures.

As Flugel has observed, the result was "the most successful neurological theory that had hitherto been propounded." It offers an ingenious "cybernetic" model, complete with receptors of signals, channels for the communication of "information," servo-mechanisms, and feed-back devices—the whole well deserving the close consideration of those who like to compare the operations of the brain with feats of modern engineering.

The doctrine of hierarchical organization was more fully expounded in one of the earliest of his many books, the little *Primer of Physiological Psychology* (6). It began by describing psychology as the "systematic study of conduct." "To define it in the traditional way as the science of consciousness is," he points out, "to exclude the study of unconscious factors." Several of his British predecessors, like Sir William Hamilton at Edinburgh and W. B. Carpenter in the University of London, had already emphasized the importance of "unconscious cerebration"; McDougall's contribution was to insist on the essentially dynamic nature of all such processes.

He now went on to apply his scheme to conative functions as well as to cognitive—to the analysis of human motives as well as of perception and thought. The followers of Bentham, Bain, and Mill still clung to the associationist doctrine that human actions are determined by a rational comparison of anticipated pleasures and anticipated pains. In place of this venerable philosophical doctrine, McDougall now contended we must substitute a frankly biological standpoint, and regard mental life as a series of adaptive responses, dictated quite as much by emotional impulse as by intelligent judgment. Like other animals, man inherits certain neural mechanisms or "dispositions"; and these, it was assumed, force him to attend to certain environmental situations as they arise, to get emotionally excited as a result, and to react, vigorously and at first almost automatically, in a way more or less appropriate to the occasion. In humbler animals such "innate goal-directed tendencies" are called instincts; and in man they appear to form the foundation out of which each of us progressively develops his own personality or character.

*These took the place of "ideas" in the older expositions. The modern term "information" (as used by the communication engineer) would seem to overcome the linguistic difficulties both of McDougall and his predecessors, and at the same time lends itself readily to a statistical interpretation of the events within the brain: (cf. Burt, *Brit. J. Statist. Psychol.*, IV, pp. 31-7, and 195-9).

†The idea that, as the nervous energy sparked across a particular cell gap, a specific sensation would flash out implied the very type of sensory atomism which later he so strongly attacked. Accordingly in later years he preferred to imagine "an everchanging field of forces," set up by the network of nerve-currents, as the nearest approach to a "seat" of consciousness. And, as we shall see, this earlier scheme underwent considerable modification in the light of his further speculations and a more detailed knowledge of the electro-chemical processes occurring in the brain.

A few years later, in a volume written for the Home University and entitled *Psychology, the Study of Behaviour*, these views were set forth in less technical terms, and applied to the study of childhood, society, and mental abnormality. This little book has been reprinted twenty-four times, and a fresh edition was called for only a couple of years ago. For the general reader it still provides the best introduction to the science.

III.—COGNITIVE PSYCHOLOGY: CHILD STUDY AND MENTAL TESTS.

When opening his new psychological laboratory in London, Sully had expressed the hope that it would "differ from those on the continent in being devoted to the study of the individual personality rather than of mind in the abstract"; and, partly with this aim in view, he induced the College authorities to grant Galton a room for his work on mental tests and body measurements. Sully himself had already founded an "Association for Child Study," and started a campaign for "child guidance." In McDougall's view these subjects were all admirably suited for experimental research. As "guinea-pigs," to borrow his own expression, there were the boys who attended University College School (then housed in what is now the Department of Psychology), and the children who formed the demonstration-classes for the Department of Education (which was also under Sully's wing), and a year or two later there were his own small infants. He started an annual course of lectures on "The Psychology of Childhood"; and, although they were unfortunately never completed for publication as he originally intended, the influence that he exercised on educational psychology, both through his pupils and through his own writings, must be counted as one of his most valuable contributions.

At University College, he came into close touch with Galton and Karl Pearson; and as a result became increasingly interested in mental measurement and mental inheritance, and an ardent supporter of eugenics (II, p.vii). One of Galton's most cherished projects was an anthropometric survey, to be carried out by a committee of the British Association and include mental measurements as well as physical. McDougall was appointed secretary; and in his laboratory at Oxford now set himself, with the aid of his research students, to devise and standardize tests of intellectual ability that could be used for such inquiries. Galton had suggested that the efficiency of the various tests might first be checked by the correlational methods which he and Pearson had devised; and accordingly McDougall invited Pearson to come to Oxford and give a brief exposition of these novel procedures. Though he deplored that he himself could scarcely calculate a mean correctly, he warmly encouraged his disciples—Brown, English, and myself—in their endeavours to adapt the Galton-Pearson techniques for the statistical analysis of mental abilities. It would, I think, be no exaggeration to say that the basic idea of factor analysis was due to McDougall.

About the same time, Spearman, a young army officer, who, with J. M. Cattell, had studied under Wundt at Leipzig, and, like Cattell, had become fascinated by the new work on mental tests, eagerly came forward to seek McDougall's advice. On returning to England he settled for a while in a Berkshire village not far from McDougall's house, and, at his suggestion, began to apply to children in the rural schools some of the tests that had been used with primitive tribes in Borneo. The results of these various investigations led to frequent and animated discussions in McDougall's garden about the broader inferences that might be drawn. Galton had recognized two types of ability—"general" and "specific." Brown held the assumption of "general ability" to be superfluous; Spearman preferred to reject the "special abilities" on the ground that these were mere relics of the obsolete

theory of faculties. McDougall held that Galton's view might be preserved if, accepting an evolutionary interpretation, we supposed that the more recent and more specialized abilities had been "formed by gradual differentiation* out of pre-existing abilities" and that the differentiation was renewed during the growth and maturation of each individual child. As he stated in one of his later books, "Spearman's theory (which seeks to dispense with so-called 'group factors') seems far too simple to account for the facts of correlation"; the mistake of both Spearman and Brown, he thought, was that of so many statistical psychologists: they trusted too exclusively to mere mathematical analysis; "the mathematical method should follow rather than lead the attack" (15, pp. 85-96, 9, pp. 66f.).

Although McDougall and Spearman both agreed in recognizing a general factor of "intelligence," there was a further difference between them in regard to its mode of working. Spearman, in those days a staunch follower of the associationist school, held that "there was nothing in the intellect which did not reach it through the senses"; and therefore argued, with Bain, Sully, and other representatives of the same tradition, that intelligence consisted essentially of a general capacity for sensory discrimination. McDougall, on the other hand, identified intelligence with the capacity for organization or integration, i.e., as he preferred to say (borrowing Stout's somewhat technical language) with "a capacity for noetic synthesis." He therefore maintained that the best tests of intelligence would be found, not in the tests of sensory discrimination that Spearman advocated, but rather in tests of apperceptive processes. Recent work has fully borne out his novel suggestions. He himself devised several ingenious procedures based on this principle; and, though the word "apperception" fell for a while into disfavour, it is encouraging to find that the usefulness of such tests has since been confirmed and extended, particularly by many American investigators—though they do not always recognize their indebtedness to McDougall.

IV.—CONATIVE PSYCHOLOGY: INSTINCTS, SENTIMENTS, AND SOCIAL LIFE.

From these investigations of cognitive tests, McDougall turned to the systematic study of conative processes. Here the same characteristic features reappear, namely, a profound conviction that the mechanistic theories of the old associationist school, as developed by Bentham and the Mills and revived by Watson and the behaviourists, are utterly inadequate to account for the phenomena of mental life, and a firm belief that the clue to most psychological problems is to be sought by starting from an evolutionary standpoint.

In this field his most permanent contribution is contained in what is undoubtedly his best known work—*An Introduction to Social Psychology*. Its primary aim was to insist that the basic science on which all the social sciences—ethics, economics, sociology, the theory of politics, and cultural anthropology—must rest is psychology. And with that end in view he sought to interpret the behaviour of man, regarded as a social animal, in terms of dynamic concepts which should do full justice to its purposive nature. While insisting that the ultimate source of all mental energy is to be traced to a limited number of innate biological impulses, he laid equal stress on the effects of personal

*The theory of the progressive differentiation of cognitive abilities appears to be fully confirmed by factor analysis (see this JOURNAL, xxiv, 1954, pp. 76-90), though most factorists have ignored it. It would appear to have special implications for the theory of mechanical models. McDougall thought it difficult to demonstrate learning in an unconscious machine. Certainly, as a little calculation will show, in a statistical mechanism which has anything like the variety of choice possessed by the adult mind, the learning, even of a comparatively simple reaction, would take years or even centuries to master. Yet, if the machine were so constituted as to "grow" by slow differentiation and integration, this theoretical difficulty might more readily be overcome. At any rate, we here apparently have one obvious explanation for the prolonged immaturity of the more intelligent creatures.

experience and environmental influences—an essential complement in his teaching which his critics have all too often overlooked.

The first half of Section I (the theoretical section of the book) is devoted to an account of the innate foundations of social life. Galton's researches on heredity had been mainly confined to intellectual abilities; and the inheritance of emotional and temperamental qualities had been largely ignored. In McDougall's view, as we have seen, their inborn basis consisted of nervous mechanisms underlying the primitive instincts which man inherits in common with other higher mammals—a hypothesis suggested by a paragraph in Darwin's *Descent of Man*. Moreover, he thought it possible to show that "each instinct conditions some specific kind of emotional excitement, which yields an access of further power to reinforce the ensuing action."

The second and larger half of the theoretical section discusses the general development of acquired motives in the growing child, and more particularly of the "self-regarding sentiment" or "ego-ideal," which, he believed, furnished the key to the whole problem of volition. Here he attempts to explain how the behaviour of any individual from infancy onwards is increasingly determined by the progressive organization of the simpler conative tendencies into habitual systems of interest and action, under the influence of the social environment. These "sentiments" (as he and other writers have termed them) are largely moulded by the cultural patterns that have been handed down, not by heredity, but by tradition and imitation within the groups to which he belongs, and then modified by the more specific attitudes of his own parents, his relatives, his school, and the social class of which he is a member. Often—indeed almost always during the earlier years—such complex emotional interests are acquired unconsciously and operate unconsciously. Thus conduct which on the surface may appear irrational, abnormal, or even anti-social, can be explained as the effect of conflicting or co-operating conative impulses of which the individual is himself unaware—a view that McDougall had already put forward in *Mind* as far back as 1898 and had briefly developed in his *Primer* and other publications.* The concluding section of the book is concerned with showing how in general these various purposive tendencies function in the daily life of both savage and civilized communities.

As a well-known American writer, Professor Gardner Murphy, has said, "this book marked the beginning of a new era in social psychology: it swept everything before it" (19, p.405). During the next twenty years over a hundred thousand copies were sold; and at the author's death it had reached the twenty-fourth edition. Nevertheless, from 1918 onwards the very notion of instinct came under heavy fire, particularly from the leaders of the behaviouristic school. So crushing were the onslaughts that, twenty years later, it was widely believed that "the doctrine of instincts is not only dead but damned."† "The result," as Gardner Murphy observes, "was to leave social psychology without any generally accepted theoretical basis" (19, p. 406). And the same unhappy consequence befell the psychological study of personality. McDougall's scheme of innate tendencies, both specific and general, and of acquired habits, attitudes, and sentiments, all forming a "hierarchical structure"‡ and crowned by a dominating "ego-ideal," had offered a most serviceable working schedule

*McDougall not unnaturally took the keenest interest in Freud's doctrine of unconscious motivation when it later became known, and was one of the first to draw general attention to it. Later he had his own dreams analysed by Jung (13, p.xi). To him and his followers much in the new psychoanalytic teaching, stripped of its paradoxical or sensational embellishments, seemed to fit in with, and even to have been anticipated by, the theories of McDougall himself (15, p.27).

†See the Symposium in this JOURNAL, "Is the Doctrine of Instincts Dead?" (XI, 1941, pp. 155-172 and following issues) and the references there given.

‡Cf. 7, pp. 431-443. This is the structure to which the modern designer of "human automata" has been led. (On the need and the possibility of constructing "a hierarchical ordering of goals in a goal-directed artefact," see D. M. MacKay, "Mentality in Machines," *Proc. Aristot. Soc. Supp.*, 1952, pp. 67f.).

for the study of individual character and the compilation of case-reports. Today every writer on personality—whether psychologist or psychiatrist, educationist or sociologist—is left to improvise his own private scheme for assessing his cases, based not on any accepted scientific terminology, but on loose popular notions all vaguely defined.

However, during the last few years the pendulum has started to swing forward once again. Under such innocent aliases as "drives," "needs," "urges," or the like, the outlawed concept has begun to show its face in several unexpected quarters; and, to quote the same historian once again, "with the passing of the intense wave of anti-McDougall feeling, instinctive mechanisms have now come slowly back into place" (19, p.443). The new American slogan—"purposive behaviourism"—would be an excellent title for what McDougall called a "hormic psychology."

In 1920 the *Introduction* was followed by *The Group Mind*. This endeavoured to formulate, on the basis of the previous work, the main "principles of collective psychology," and apply them to the interpretation of national life and character. Here the "culture-pattern"* of the larger and more complex type of group is itself described as having a loose "hierarchical" organization; every constituent group—whether nation, tribe, or social class—is said to possess its own collective opinions and ideals, its own "group spirit," which, working primarily through the family and other smaller units, exerts a strong and steady pressure on the growing child and the full-grown adult, so as to shape the personality and conduct of each member more or less in conformity with the same implicit ideals. McDougall's views were derived partly from his own observations among primitive tribes, and partly from the sociological theories advanced by Hobhouse, to whom the book is dedicated. They exercised considerable influence on younger anthropologists like Malinowski, who in his turn has influenced such eminent authorities as Ruth Benedict and Margaret Mead.†

V.—BODY AND MIND.

Someone has rather caustically observed that the publication and re-publication of the *Introduction to Social Psychology* was rather like the packing and re-packing of a portmanteau for a journey that was never undertaken. The change of plan was due to an unforeseen invitation. In 1920 McDougall accepted the Chair of Psychology at Harvard—a post that had been held by James and enjoyed the highest prestige in the psychological world. There he found himself beset by a hostile band of behaviourists, who were proclaiming a crude mechanistic version of the old physiological associationism, long defunct in England. He therefore deemed it his duty to embark on a series of polemical and semi-popular publications, "exposing the preposterous absurdity of a psychology without a soul," and developing and defending what he now began to call his "hormic theory"—a theory which maintained that purpose itself could operate as a *vera causa* and that its operation was now an established fact.

To many this appeared a sad relapse from the "physiological psychology" of his early *Primer*. There he had started with two cardinal premisses based

*The term culture is used in the sense defined by Tylor—"that complex whole which includes belief, knowledge, art, law, custom, morals, and other generalized habits and ideas acquired by a man as a member of his society."

†The chief difference is that, whereas the latter imbibed their psychology mainly from the behaviouristic school, and consequently sought to explain almost all social behaviour by social conditioning, McDougall strove to give full weight to both nature and nurture.

on the latest neurological research. First, as experiments on decapitated animals had plainly shown, "the simpler reflexes of the spinal cord seem to be completely determined by a chain of purely physiological processes." Secondly, as had been indicated by the recent work of Sherrington and his contemporaries, "the structure and functioning of the rest of the central nervous system (the 'brain') is essentially similar to that of the spinal cord: it consists of nervous arcs of great complexity serving essentially to transmit physical excitations from sensory receptors to muscular effectors." By ordinary syllogistic reasoning it would seem to follow that the whole behaviour of man is therefore determined in the same reflex fashion as the mechanical activities of the spinal cord: we are in fact physiological machines, and "there is no room for the intervention of mind as a directing agency."

McDougall, however, deliberately stopped short of this conclusion. Warned by the mistakes of his predecessors, he had a profound distrust for mere unconfirmed deductions. In natural science, as he repeatedly declared, the only cogent procedure is to push the simpler and more readily verifiable types of explanation as far as they will go, and then see what is still needed to account for the recalcitrant residuum. Accordingly, in his laboratory at Oxford he spent much of his time discussing, designing, and even constructing a variety of diagrammatic and working models, which might illustrate or reproduce typical activities of the human mind. In view of the renewed interest in such inventions, it will be of interest to note how far the theoretical principles on which he relied have since been adopted by contemporary mechanicians.

The earliest attempts at constructing a mechanical man were those of the Swiss watch-makers. Their device was to insert a clockwork mechanism into a life-size doll, which would then jerk its legs, wave its arms, or grind out a couple of tunes, like Olympia in the *Tales of Hoffman*, but of course display no adaptive responses to outer stimuli. Descartes, impressed, it is said, by the ingenious water-works that were being planned at Amsterdam, preferred to formulate his theory of "reflex action" in terms of a hydraulic rather than a mechanical model: a current of "animal spirits" was assumed to flow along the tubular nerves, and so activate the muscles. Frankenstein, the medical student in Mary Shelley's novel, constructed his monster from the bones and muscles of dead animals, and used a current from a voltaic pile to animate the limbs.* One uniform characteristic of all such contrivances, both old and new, and of the speculations that give rise to them, is the constant endeavour to reinterpret mental behaviour in terms of the latest sensational achievement of popular science.

McDougall's earliest models, like those of Descartes, were conceived mainly in hydraulic terms: the essential work was done by "currents" of nervous energy, flowing from a common "reservoir," and directed by variable sluice-gates situated at the cell-junctions or "synapses." His famous theory of "inhibition by drainage" was illustrated by tubes and taps that functioned like a syphon. In his lectures, however, he usually began by comparing and contrasting the human brain with an imaginary self-operating telephone-exchange. An improved theoretical analogue was then built up by successively introducing half-a-dozen modifications. First, he maintained that each "wave" of nervous energy must be cumulative, "more like firing a train of gunpowder than conducting an electric current"—a view hotly disputed at the time, but since fully confirmed. Secondly, he claimed that the phenomena of apperception clearly proved that the signals transmitted were not isolated currents, as the stimulus-response hypothesis of the earlier experimentalists implied, but complex messages conveyed as patterns. Thirdly, he pointed out that some

*From her preface to the later editions it appears that the problem discussed was inspired by Shelley's philosophical reflections on his own experiences as a medical student and his experiments in his laboratory at Field Place. Minor details, she adds, were also drawn from "experiments ascribed to Dr. (Erasmus) Darwin and certain German physiologists."

of the stimuli determining the temporal course of any skilled movement were themselves the direct or indirect effect of the preceding phases of the same movement and resulted from the incidental excitation of "proprio-ceptive" organs. This was suggested by Sherrington's recent work on the newly discovered muscle-sense, which guides and governs our posture and our actions. A similar "looptine mechanism" was assumed in his theory of the visceral reinforcement of emotional responses. In developing these ideas, it will be noted, he was one of the earliest to emphasize the function in the nervous system of what the modern engineer would call the principle of "feed-back"; and two of his favourite analogies were drawn from the ingenious inventions of James Watt—namely, the safety-valve and the automatic governor. Fourthly, and most important of all, he maintained that transmission across the synapses was not like that at an ordinary electrical switch, which is either "on" or "off," but (as Fechner's experiments on discrimination suggested) a variable process which could be only expressed by probabilities: the ultimate discharge might be of an all-or-none nature, but the raising or lowering of resistance above or below the critical "threshold" was itself a matter of degree. Here then was an inescapable element of indeterminacy. This (he argued) is most plainly shown by experiments on mental tests and on habit formation: they demonstrate that, at the very best, all that an external observer can ever do is to predict the *probable* performance of the reagent. As he remarked in discussing learning by trial and error, "the earlier reactions of the higher cerebral levels have often a random character, as though the choice was determined by blindly throwing dice: the problem therefore is—can consciousness increase efficiency by loading the dice?"* All blueprints for an artificial brain should, he concluded, be based in future not on mechanistic, but on "probabilistic" principles: the statistical models of Galton, where the detailed movements were unpredictable, supplied far better parallels for the working of the mind than the calculating machines of Babbage or the "logic machine" of Jevons, whose perfect accuracy "showed how inhuman they were." And for much the same reasons he was greatly attracted by Thomson's "sampling theory" of mental abilities.†

The outcome of all these speculative efforts was that he grew more and more firmly convinced that, as he had vaguely suspected from the outset, there would for ever remain certain irreducible characteristics in human behaviour which would defy all attempts at mechanistic explanation. In his later volumes he set himself to specify them more precisely. On the cognitive side he pointed to various forms of memory. James had suggested that the "power to learn" distinguished the brain from the spinal cord, and formed a distinctive mark of a conscious mind.‡ The older types of artificial machine could only do what they were programmed to do: no musical box can spontaneously learn a new tune; no clock can teach itself to tell the time more correctly. However, the invention of the thermionic valve and kindred devices has greatly enlarged the possibilities open to the modern gadgeteer. Of more recent attempts to construct a mind-like apparatus, one of the earliest—that of Thomas Ross in 1938—actually succeeded in producing a machine which, like the laboratory rat, could learn to find its way out of a maze by repeated trial and error. McDougall admitted that "habit formation" might be imitated in an artificial robot; and under habit he included "memory resulting from sheer repetition." But "pure memory"—e.g., feats of recognition, based on a single experience only—must, he believed, involve an entirely different type of

*12 p. 19, 3f. Cf. also James, *Principles*, I, p. 140.

†Except on paper, models incorporating statistical or stochastic processes are by no means easy to construct. But among post-war machines, I suspect, he would unquestionably have preferred the principles of the homeostat, described by Ashby in his *Design for a Brain* (1952), to those of the so-called "electronic brains" which have so greatly impressed the popular writer.

‡W. James, *Principles*, I, pp. 20f.; also W. McDougall, *The Energies of Men*, ch. xxiii, "Learning and the Steering Process."

activity—psychical rather than physiological. "Habit is of the body, memory of the mind."* Curiously enough, as it has since turned out, recognition is in fact a process that has proved exceedingly difficult to imitate in a lifeless contrivance.

On the conative side he laid chief stress on purpose—the keynote of all his later writings. Nor did he ever tire of sharpening the antithesis between mechanism and teleology (14, 16, 17). The first characteristic of conation is the tendency towards readjustment. This is a conspicuous feature of all living things; but it is also observable at a lowlier level. Stout, in searching for "a mode of physical self-determination" which might form the counterpart of conative activity, found it in "the process whereby a neural system regains equilibrium after disturbance" (2, I, pp. 148f. and 268). Ashby's homeostat would fit Stout's specifications admirably. The thermostat in the kitchen refrigerator affords a still more familiar instance. But McDougall insists that Stout's requirements are insufficient. The object of purposive striving is seldom a return to the *status quo ante*: there is usually some unattained end or goal towards which the creature directs his efforts. In his view, therefore, conation has a second characteristic—self-direction. Psychic energy, or "hormé" as he preferred to call it,† is "directed energy"—a concept that he never succeeded in making quite clear. At its simplest perhaps goal-seeking might be defined as an activity that tends to minimize the distance between the pursuer and the object (or state) pursued, no matter how that object shifts and changes during the process. The self-guided missiles developed during the war would seem to fulfil this further criterion without invoking any novel brand of energy. McDougall, however, adds a third criterion—the "selectivity of purpose." Human beings not only seek their goal; they choose it. Choice is by no means an undetermined act of "free will": it is an act depending on the recognition of values, and this in turn depends on consciousness. After all, it is consciousness that distinguishes mind from matter; and no engineer, however skilful, can infuse consciousness into his machine. The behaviourist, of course, retorts that consciousness is superfluous: but McDougall counters this by asking—would any of them be satisfied with an automatic sweetheart or with mindless friends? It is just because each of us is a *conscious* personality that we value a human being above all else. When a father kills his offspring, everyone holds him guilty of the gravest of crimes; but no one would want to hang Dr. Grey Walter just because he demolished one of his own "mindlike models."

The "hormic theory," as thus developed, forms a recurrent theme in each of McDougall's later books. But the most systematic exposition is to be found in a weighty tome entitled *Body and Mind*, which he always regarded as his *magnum opus* (8). Here he attempts a thorough examination of all the principal hypotheses that have ever been put forward to describe the relation between brain and consciousness. As he said, when presenting me with a copy fresh from the press and patting its covers with a smile like that of the satisfied tiger: "I believe I've got all of them inside." The main argument proceeds by the

*Quoted by McDougall from Bergson, *Matière et Mémoire*. To explain recognition the favourite device of post-war designers is to exploit the principle of "scanning." My own preference is to express all apperceptive processes in terms of factorial concepts—by matrices (to designate patterns) and correlations (to designate likelihoods). See *Psycho-metrika*, III, p. 160.

†The term was borrowed from Sir Percy Nunn: "hormé (purposive drive or urge) is the basis of those activities that differentiate the living animal from dead matter" (*Education: Its Data and First Principles*, p. 28). Note that, as thus defined, the term covers vital as well as psychic activities.

elimination of each rival theory in turn, until only one is left which will, with any plausibility, account for all the facts; and this is a renovated version of the old interactionist theory, which postulates a psychic principle or medium capable of interacting causally with the physiological mechanism or "brain."

He was, however, by no means content to set forth the pros and cons in terms of armchair analogies and inferences. He held that the view to which he had been led was both suggested by, and needed to be further confirmed by, direct experimental evidence. In his early work on vision, he had shown that experiments on colour-mixture and the binocular perception of space revealed modes of combination quite unlike those of mechanical or chemical processes—e.g., the fusion of red, green, and blue, which yields, not reddish greenish blue but white, and the fusion of two retinal pictures, which produces a single stereoscopic percept characterized by the quality of depth. In the future, he thought, certain crucial experiments might be achieved by operating on the living brain. "If ever I am seized by an incurable illness," he announced, "I have asked Sherrington to cut through my corpus callosum with his scalpel, and so separate my cerebral hemispheres; I hope I shall then be able to say whether I am a single personality or two conscious streams."*

Surprisingly enough, such experiments have since been carried out. "Perhaps the most astonishing of all observations in the field of cerebral surgery," writes Dr. Wilkie, "are those made on human beings in whom the whole corpus callosum and anterior commissure have been severed, thus cutting all known direct communications between the left and right hemispheres." One or other of two conceivable consequences might be anticipated as a result of this extraordinary operation: either the patient would report distinct but concurrent activities running separately in the two hemispheres, or one hemisphere might be expected to dominate, while the activities of the other would be suppressed. Contrary to all expectation, no such after-effects have been found. "Binocular fusion and the binocular perception of depth," we are told, "are not materially altered." "The patient suffers almost no inconvenience (*sic.*) Changes in his behaviour can be detected only by refined techniques, and these turn out to be rather trivial, consisting chiefly in a lengthening of the reaction time."†

His later books, it must be confessed, are somewhat disappointing. In arguing with his American critics, he was apt to write in a somewhat pretentious style, which one of them described as "Scottish baronial." On the theoretical side, the only fresh contribution of any importance is the insistence on the view that life and consciousness are continuous, both being manifestations of the same hormic energy (14). For a while he seemed to be groping towards a hypothesis of "emergent evolution," similar to that so eloquently expounded by Lloyd Morgan and Professor Broad. In the end, however, he decided that "both the words that make up this fashionable phrase are question begging terms"; and finally pronounced for a "frank and uncompromising bifurcation" (14, 17). But, like Bergson and Driesch, he preferred to draw the main line of separation, not between unconscious matter and conscious mind, but between inanimate objects and living organisms. Previously most dualists had held that the real irreducible contrast lies between mechanical processes, on the one hand, and the inescapable fact of conscious awareness, on the other—a contrast which (in Stout's phrase) "constitutes an unmediated and inexplic-

* Cf this *Journal*, IX, 1939, p. 2.

† For a more detailed summary of the results, with references to the neurological literature, see J. S. Wilkie, *The Science of Mind and Brain*, 1953, pp. 133f. For reasons that would here be out of place, I should not myself regard such observations as decisive in themselves. In psychology no single experiment, however impressive, can be absolutely conclusive.

able breach in the unity and continuity of the universe" (*Mind and Matter*, p. 66). However, here we are encroaching on the realm of metaphysics; here psychology is, or should be, silent, though no doubt psychologists are, and will continue to be, voluble.

McDougall's biological speculations led him in his last few years to re-investigate the old Lamarckian hypothesis which he had so vigorously criticized in his earlier writings on eugenics. With the help of his son, he carried out an elaborate series of experiments on forty-nine successive generations of rats. The animals were regularly trained to learn a difficult task by trial and error, and the outcome seemed to be that "later generations showed an increased facility, measured by a reduction in the amount of training required by nearly 90 per cent." In his opinion this "inheritance of an acquired characteristic" went far to confirm the view that evolution is no mere mechanical process but itself essentially purposive, and "lent strong support to the idea of a non-material basis for racial as well as individual memory." Few, I fancy, would nowadays consider this a conclusive or a permanent contribution. But the experiments certainly deserve to be repeated.

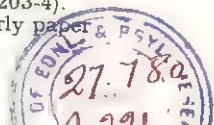
Far more influential was his work in the field of psychical research. His early experiments on hypnotism, carried out with his students at Oxford, were considered "most unprofessional" by his medical colleagues, and gave rise to a horrified rumour among the clergy that he was "reviving the Black Art."* Only the intervention of Osler, the Regius Professor of Medicine and the foremost physician of his day, saved him from forfeiting his post. But his demonstrations successfully convinced most British psychologists that "in spite of continued scepticism, hypnosis is a genuinely useful and easily produced phenomenon"; and the theory of sleep, suggestion, and hypnotic trance to which he was led, and the widespread interest which was thus aroused, had far-reaching effects on the views of later writers.†

In 1927 he moved to Duke University in North Carolina. Hoping to secure more direct evidence for a purely psychic principle, he now planned a number of systematic investigations on the problems of "paranormal psychology"—mediumship, clairvoyance, and above all telepathy—topics that were still severely frowned upon by most reputable scientists. One important outcome was the remarkable series of experiments by his colleague, Dr. J. B. Rhine, on "extra-sensory perception," which in McDougall's view "placed the hypothesis of telepathic communication on a firm experimental basis." His bold decision to prosecute such researches in a University laboratory has of itself had a revolutionary influence on the attitude of the academic world. In his own former laboratory at University College, Soal and others have since attacked the problem by similar methods with improved statistical techniques, and doctorates have even been awarded for theses in the field of psychical research. Still more recently at his old Universities of Oxford and Cambridge, fellowships have been instituted for the systematic study of the subjects that once were so rigorously tabooed.

His last publication consisted in a critical survey of all the current forms of materialistic or "mechanical biology." It included a brief account of the experiments just described, and ended with a plea for "frankly admitting the

*Less than a century before, John Elliotson, Professor of Medicine at University College—the Dr. Goodenough of Thackeray's *Pendennis*—had experimented with hypnotism as a cure for hysteria and given demonstrations (attended by Charles Dickens) in the hospital theatre. He too had been compelled to resign his chair: (*Lancet*, II, 1838, pp. 203-4).

†Cf. his article on "Hypnotism" in the *Encyclopaedia Britannica*, and his early paper on "The State of the Brain during Hypnosis," *Brain*, XXXI, 1908, pp. 67-81.



efficacy of psychical activities, as teleological guiding influences." The book was entitled *The Riddle of Life*, a riddle which he tried so long and so hard to solve; and a few months after it was completed he died at the comparatively early age of 67.

So much that McDougall contended for is now taken tacitly for granted, that few of the younger generation, either in this country or America, realize how profound and widespread was his influence. When he started his first experiments, a materialistic version of the old associationist doctrines, based on a ludicrously simplified form of sensory atomism, still dominated the teaching of nearly all his fellow-pioneers, who were then struggling to convert psychology from a branch of metaphysics into a branch of natural science. Philosophers like Ward and Stout, it is true, had already begun to register their a priori protests. But it was McDougall, by his ingenious experiments, his lucid and readable style, and his prolific output of technical and popular publications, who did most to effect the change. No other book in the whole range of psychology ever approached the sales of his two best sellers; no other publications of that date are still so widely read by students. The seeds of his ideas can doubtless be traced, buried with others less fruitful, in the writings of his predecessors. He stands out from all the rest by his genius for seizing just those particular points that deserve to be tested or developed, and by his unrivalled skill for organizing the results into a coherent and comprehensive theory. As Flugel observes, "many details of his system will unquestionably need modification or restatement as our knowledge grows; but McDougall's attempt at an orderly presentation would still seem to be more hopeful than any other in providing a scaffolding by means of which a more firmly established edifice may eventually be built."

REFERENCES

1. SULLY, J. (1892): *The Human Mind: A Textbook in Psychology* (Longmans, Green and Co.).
2. STOUT, G. F. (1896): *Analytic Psychology* (Sonnenschein).
3. MCDUGALL, W. and HOSE, C. (1900): *The Pagan Tribes of Borneo* (Cambridge University Press).
4. MCDUGALL, W. (1901): "Some New Observations in Support of Thomas Young's Theory of Light and Colour-Vision," *Mind*, N.S. X, pp. 52-97, 210-245, 347-382.
5. ——— (1902): "Physiological Factors in the Attention Process," *Mind*, N.S. XI, pp. 109-132, 216-229, 324-340.
6. ——— (1905): *Primer of Physiological Psychology* (Dent).
7. ——— (1908): *An Introduction to Social Psychology* (Methuen).
8. ——— (1911): *Body and Mind* (Methuen).
9. ——— (1912): *Psychology: The Study of Behaviour* (Oxford University Press).
10. ——— (1920): *The Group Mind* (Cambridge University Press).
11. ——— (1921): *National Welfare and National Decay* (Methuen).
12. ——— (1923): *An Outline of Psychology* (Methuen).
13. ——— (1926): *An Outline of Abnormal Psychology* (Methuen).
14. ——— (1929): *Modern Materialism and Emergent Evolution* (Methuen).
15. ——— (1932): *The Energies of Men: A Study of the Fundamentals of Dynamic Psychology* (Methuen).
16. ——— (1934): *The Frontiers of Psychology* (Methuen).
17. ——— (1938): *The Riddle of Life* (Methuen).
18. FLUGEL, J. C. (1933): *A Hundred Years of Psychology* (Duckworth).
19. MURPHY, G. (1949): *Historical Introduction to Modern Psychology* (Routledge and Kegan Paul).

THE CAUSES OF PREMATURE LEAVING FROM GRAMMAR SCHOOLS

PART II

BY MILDRED COLLINS

(Lecturer in Education, University College, Leicester)

I.—*Aims of Investigation*: (i) *Selection procedures at 11 plus and premature leaving.* (ii) *Provision of secondary technical education and the reduction of premature leavers.* (iii) *"Late-transference" between grammar and secondary modern schools and prevention of premature leaving from the former.* (iv) *Four-year courses to G.C.E. at ordinary level and the encouragement of early leaving.* II.—*General Conclusions.* III.—*Summary.* IV.—*Appendix.*¹

AIMS OF INVESTIGATION

In the first² of these two articles on premature leaving from grammar schools in England, it was shown that withdrawal from these schools before reaching 16 years of age is more common in those areas with higher proportions of unskilled workers³, higher percentages of grammar school entrants, and with more difficulty in filling teaching posts in the grammar schools. There was evidence also that the possibility of a "fine" for premature withdrawal discouraged this practice. But the associations were not close and it is possible that more efficient methods of selection for, and greater variety within, the secondary educational system of some areas may play important parts. This article therefore deals with the more common selection procedures, the provision of secondary technical education, the "late" transference from one type of secondary school to another, and the provision of four-year G.C.E. courses in some grammar schools. The information has been obtained largely by means of questionnaires to L.E.As. Not all the 127 English authorities replied, and of those which did, not all supplied sufficient information for use in analysis, so that the results are based on a sample, which, however, does not seem to be unduly biased.

(i) *Selection procedures at 11 plus and premature leaving.*

Sixty-three of the 127 English L.E.As. supplied information about their methods of selection—or allocation—for secondary schools as applied in 1947 and 1948, when presumably most of the premature leavers of 1950, 1951 and 1952 (the period under review here) were admitted. These sixty-three authorities form a fairly representative sample as far as the criteria discussed in the

¹As for the first of these two articles, I should like to thank those L.E.As. who have supplied information for this enquiry, the Ministry of Education for providing certain figures and Miss F. Conway for help in the statistical treatment of some of the data. I alone am responsible for the selection of the form in which the information is presented and any conclusions drawn from it.

²Collins, M. "Causes of Premature Leaving from Grammar Schools," this JOURNAL (1954), XXIV. pp. 129-142.

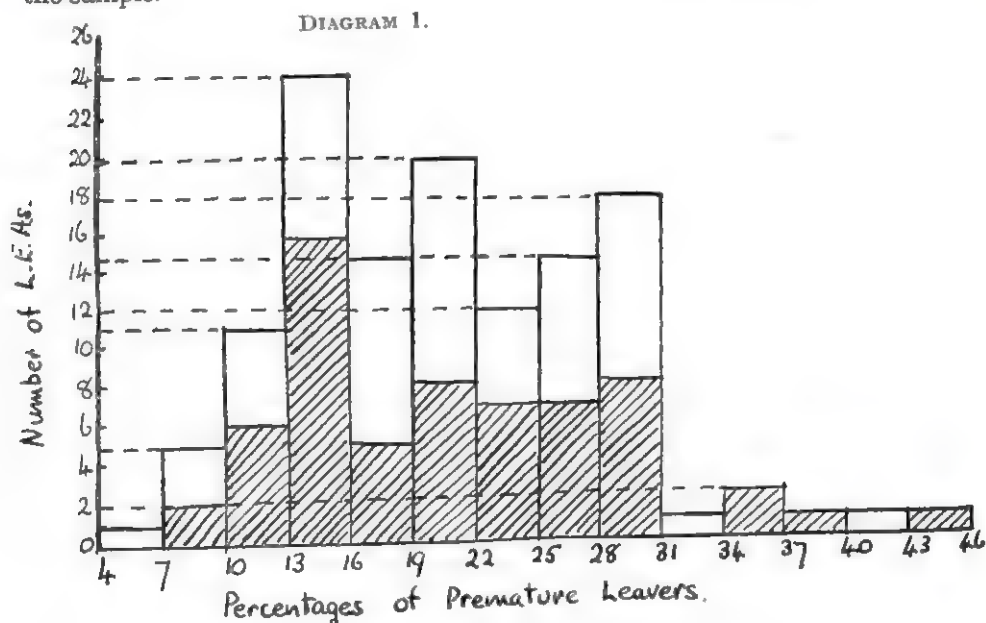
³An estimate of the proportions of social classes has been obtained from the 1 per cent. sample of the 1951 Census figures (see *Census, 1951, Great Britain—One per cent. sample Tables, Part I* (London: Her Majesty's Stationery Office.) pp. 83-85.) Classes IV and V of the Census are grouped together as "unskilled workers" in this investigation and Classes I and II are put together and called "professional workers," although all these classes contain workers who cannot be aptly described by these terms. The percentages refer to males over 15 years of age.

first article are concerned. Their distribution from the point of view of percentage of premature leavers can be seen by consulting Diagram 1. The average of their percentages of grammar school places at 13+ in 1950 was 20.2 per cent. (S.D.=5.41) compared with 19.9 per cent. (S.D.=5.82) for the country as a whole. The sample contains forty-one out of the seventy-nine County Boroughs, and when the L.E.As. are allocated to the appropriate groups according to increased percentages of premature leavers—after allowances have been made for the correlation with percentages of entrants¹—then there are thirty in groups I and II, the better groups, with relatively fewer leavers, and thirty-three in the poorer, III and IV. The mean percentages of social classes IV and V in these four groups of sample L.E.As. are as follows: I=25.6, II=29.6, III=31.2, and IV=30.5.

PREMATURE LEAVERS FROM GRAMMAR SCHOOLS

1949-1952. The P.Ls. (premature leavers) are expressed as percentages of all leavers from maintained grammar schools and the average has been found for the three academic years under review.

The shaded part indicates the distribution of the L.E.As. represented in the sample.



Mean for whole of England 1949-1950=23.3 per cent.

" " " " " 1950-1951=20.9 per cent.

" " " " " 1951-1952=18.6 per cent.

" " " " " 1949-1952=21.9 per cent.

" " sample of L.E.As. 1949-1952=20.2 per cent.

The replies to the questionnaires returned by the sample L.E.As. give an impression of a keen interest in making selection as effective as possible. Since

¹For details of the method used for this classification see Collins, M. "Causes of Premature Leaving from Grammar schools" this JOURNAL (1954) xxiv p. 135. Essentially the method consists of ranking the L.E.As. according to the differences in their positions for entrants and premature leavers compared with the rest of the L.E.As.

1948, forty-one of the sixty-three report some change in procedure: the chief trends being the introduction of standardised attainment tests—very few authorities now being without them; a greater stress on school reports with the use of school record cards; experimentation with an essay of some kind; and the holding of part of the examination in the schools to which the candidates belong. The purpose here is to see whether there is any evidence that certain of these selective procedures are helpful in reducing premature leaving. An analysis of methods used in 1947 and 1948 is given below and the sixty-three authorities are put in groups I to IV according to the degree of premature leaving but taking into consideration also the percentage of grammar school entrants. If any particular technique, not uniformly used, is markedly helpful then there should be some indication of this in its distribution among the different groups.

(a) *Standardised Attainment Tests in the selection examination.*

By 1947, fifty of the sixty-three authorities were using such tests, and of the thirteen which were not, nine reported having introduced them since 1948.

NUMBER OF L.E.As.

	Group I "Best"	Group II	Group III	Group IV "Worst"	Totals
No standardised attainment tests	3	3	4	3	13
Total in each group	16	14	16	17	63

(b) *Intelligence Tests.*

Since 1948, two of the three L.E.As. formerly without intelligence tests now have them. One authority has since dropped its intelligence test. The tests are frequently, but by no means always, those of Moray House.

NUMBER OF L.E.As.

	Group I	Group II	Group III	Group IV	Totals
No intelligence test	0	1	1	1	3
One intelligence test	7	3	3	8	21
Two or more intelligence tests	9	10	12	8	39
Totals	16	14	16	17	63

(c) *School Reports and Assessment.*

In nine cases nothing of this kind was required; of the remaining fifty-four, seven demanded reports for borderline cases only. Six L.E.As. mention either having introduced school assessments or having begun to pay more attention to them since 1948. Unfortunately no discrimination was made between the different kinds of school reports required.

NUMBER OF L.E.As.

	Group I	Group II	Group III	Group IV	Total
Without school reports on candidates.....	2	3	1	3	9

(d) *Essays.*

By the end of 1948, 35 of the sample L.E.As. included an essay in their selection procedure and four other authorities have introduced an essay since. In one additional case it was used for borderline candidates only. In some forms of selection the essay was included in the general English paper. In some its mark or grade was kept separate from the rest of the marks, owing presumably to the difficulty of marking.¹

NUMBER OF L.E.As.

	Group I	Group II	Group III	Group IV	Total
No essay included in selection procedure ..	7	6	7	7	27

(e) *Interviews.*

Twenty-five authorities had no interview of any kind up to 1948, one stating categorically that it did not believe in interviews, four had interviews with the Heads of junior schools, but not with the candidates, while thirty-four had interviews with the pupils, most commonly with the borderline candidates only. No L.E.A. mentioned having introduced interviews since 1948.

(a)

NUMBER OF L.E.As.

	Group I	Group II	Group III	Group IV	Totals
No interviews with candidate	8	4	8	9	29
Neither interview nor essay	4	2	4	4	14
With interview and essay	4	7	5	5	21

The most striking feature about this analysis² is that, with perhaps two exceptions, the different kinds of selection techniques occur uniformly among

¹ VALENTINE, C. W. *The Reliability of Examinations*. (London, University of London Press, 1932), pp. 26-30.

WISEMAN, W. "Marking an Essay Reliably." *Times Educational Supplement*, 9th April, 1949.

² It is possible to select the four groups used in the analysis by a different method from the one used in this article. The "influence" of entrance can be allowed for by plotting the regression of the percentages of P.L.s. on the percentage of entrants and then cutting off appropriate areas equidistant from this line on each side. Those L.E.As. found beyond the limiting lines form Groups I and IV while those between these lines and the line of regression form Groups II and III. An analysis using this grouping is given in the appendix; the conclusions to be drawn from it do not differ significantly from those derived from the analysis discussed here. Another way of considering the data is to ignore the association of premature leaving with entrance and to divide the L.E.As. into four groups according to percentages of premature leavers only. An analysis using this method is also given in the appendix.

(b)

NUMBER OF L.E.As.

	Groups I and II "Better"		Groups III and IV "Poorer"		Totals
	Below 30 per cent. social classes IV and V	30 per cent. and above social classes IV and V	Below 30 per cent. social classes IV and V	30 per cent. and above social classes IV and V	
With interviews for candidates	10	8	6	10	34
No interview for candidates	5	7	7	10	29
Totals	15	15	13	20	63

groups I to IV. There is no evidence that selection procedures as here differentiated, help to prevent premature leaving such as was found with the imposition of penalties. The two possible exceptions are the presence of an interview and two intelligence tests, but the differences in this sample turn out not to be statistically significant¹. It should however be borne in mind that the negative results of this analysis of the varying effectiveness of different means of selection relates only to a sample of authorities and that even for this sample absence of evidence does not constitute a positive proof that evidence is not there.

A consideration of the social composition of the different groups might throw some light on the problem. When the social class distribution of the various areas is taken into account as well as a division into two main groups for premature leaving (by putting I and II together and III and IV together), it is apparent that two-thirds of the L.E.As. with below 30 per cent. of unskilled workers in the "better group" had interviews, whereas among similar L.E.As. in the "poorer" group, about half did so, but this difference is again not statistically significant.² Interviews were less common in this sample among the areas with more than 30 per cent. of unskilled workers; 51 per cent. had them, compared with 57 per cent. in the areas with fewer unskilled workers. If one considers those L.E.As. which did relatively better than expected (i.e., those in groups I and II with 30 per cent. or more of unskilled workers) and those which did relatively worse (i.e., those in III and IV with less than 30 per cent. of unskilled workers) then there is no obvious association between the inclusion of an interview in selection procedure and the reduction of premature leaving.

Actually, of course, the form the interview takes varies considerably in different areas; in some the interview is virtually another examination, in others it is a much more personal and individual affair. Similar differences, not brought out in the analysis here, also apply to the other techniques, particularly to the school assessments. Also, perhaps, it is not legitimate to isolate

¹ (i) Interview more frequent in Groups I and II, than in Groups III and IV. Chi-squared=0.830 $P > 0.40$.
(ii) One or no intelligence test more common than two intelligence tests in the "extreme" groups (I and IV) than in Groups II and III. Chi-squared=3.0056 $P > 0.05$.

² (Chi-squared=0.709 $P = > 0.40$.)

the parts of what is essentially a composite examination : the selection stands or falls by the process as a whole ; a weakness here and there may be counter-balanced by strength at other points. However, the investigation of this sample indicates that the introduction of any essay, any interview or any school record will not of itself necessarily bring about a substantial decrease in premature leaving.

Finally, it should be noted that these various techniques are being applied at different intellectual levels within the school populations. An examination adequate for selecting a comparatively small highly intelligent group may not be so suitable for picking out a larger one of lower average intelligence. And selection examinations certainly have to discriminate at different intelligence levels not only because of the variation in percentages of grammar school places available but also because of the difference in mean level of intelligence of different areas. Only very limited information is available from L.E.As. about these intelligence levels as measured by Moray House tests, and some of this information is of uncertain interpretation, but it is interesting to notice that the five authorities with lower than average I.Q. scores for the 11 plus population are all either in group III or group IV.

Groups	Group I	Group II	Group III	Group IV
L.E.As. from which information was available	2	5	3	7
Range of Mean I.Q. score.....	116* to 103	108** to 100.1	106.56** to 99.6	107.5** to 95.4

* Based on those who wished to enter for the selection examination.

** Probably also from slightly selected groups.

(ii) *Provision of Secondary Technical Schools and the reduction of premature leaving.*

In some areas the selection techniques at 11+ are used for allocating children not only to grammar schools but also to secondary technical schools. A few L.E.As. now include special space perception tests for this purpose. In other areas the older method of selection for technical education at thirteen years is still retained, while in others this type of education begins only after secondary education has been completed. It may be argued that where provision is made for entrance to technical schools at eleven or thirteen then there should be fewer premature leavers from grammar schools because an alternative for the intelligent but more "practically minded" children is available. And for those parents whose eye is kept more sharply on the "job-relatedness" of education, a technical school would seem to provide the answer more clearly than a grammar school. On the other hand in those areas where there is a large intake into grammar schools and also a liberal supply of technical secondary schools (see Table I) it is possible that there may not be a sufficient number of children able to profit from these more demanding types of education, thus encouraging premature leaving.

Nineteen of the twenty-eight L.E.As. without technical secondary education and twelve of the twenty-two with less than 2 per cent. were County authorities. No County had higher than 10 per cent.

TABLE 1

PROVISION OF TECHNICAL¹ AND GRAMMAR SCHOOL EDUCATION IN THE SAME AREAS.
(Figures based on Ministry of Education Statistical Return No. 28 for January, 1950.)

	% of Grammar School Places													No. of L.E.Hs.
	6-	9-	12-	15-	18-	21-	24-	27-	30-	33-	36-	39-	42-	45
% of Technical School Places.	26.0—27.9									1				= 1
	24.0—25.9													
	22.0—23.9													
	20.0—21.9					1								= 1
	18.0—19.9				1									= 1
	16.0—17.9	1				2								= 3
	14.0—15.9						2							= 2
	12.0—13.9					1	2							= 3
	10.0—11.9					1				1				= 2
	8.0—9.9			2	2	2		2						= 8
	6.0—7.9		1	2	2	1	2	2						= 10
	4.0—5.9	1	1	1	8	4	4	2						= 21
	2.0—3.9		2	5	2	3	6	4	2		1			= 25
	.01—1.9		1	4	2	2	5	4	1		2			= 22
	0	1		4	5	5	2	5	3		1	1	1	= 28
														127

In fact when one considers the mean percentage of technical school places of all the 127 L.E.As. grouped for leavers as before, one finds that the best group, Group I, had the same as the poorest, Group IV, in 1950.

127 L.E.As.	Groups ..	I	II	III	IV
Mean percentage of technical secondary school places		4.7	3.5	4.2	4.7
Number of L.E.As. with no provision of technical secondary education		6	7	7	6

The provision of technical secondary education does not, from this evidence, reduce premature leaving from grammar schools by providing an alternative

¹Includes Technical "streams" also.

to them. This alternative is presented however by different authorities at different times, some when children are eleven and some when they are thirteen. The former was more common among the sample L.E.As. in Groups III and IV; if anything the alternative of grammar or technical as well as secondary modern education at 11+ is associated with relatively more premature leaving, although the difference is not statistically significant¹ for the sample and information is not available for the other L.E.As. Should the association prove to be more than a chance affair then it may be due to the difficulties involved in selection for technical education at the age of² eleven or to the fact

	Groups				
	I	II	III	IV	
63 Sample L.E.As.					
No provision of technical secondary education	6	5	4	4	= 19
Provision at 11+	4	3	5	8	= 20
Provision at 13+	6	6	7	5	= 24

that any possible advantage of this additional alternative is offset by the demand for young labour in those industrial districts in which its greater provision is frequently found, and which are relatively more common in Groups III and IV.

Transference to grammar schools after eleven years of age is commonly thought to be more frequent than transference from them, and evidence from the questionnaire supports this view. Some authorities, however, transfer more from grammar schools than they do to them, as can be seen from (5) Table 2. This table gives the records of forty-nine L.E.As. for the years 1950-1951, the average of the two years being taken. The size of the school populations varies of course and the individual figures for sections 1 to 5 of the table are only useful for noticing the ratio between transference to and from grammar schools. It is clear that in this sample of L.E.As. there is a tendency for a higher proportion of the moves to be *from* grammar schools in those areas where the percentage of premature leavers is higher. The total transference, however, is almost twice as high in the L.E.As. of the "Better" groups, some of which transferred relatively large numbers to grammar schools in 1950 and 1951. This may be one indication of the keenness for that type of education in those areas. There is no evidence here that transference from grammar schools to secondary modern is associated with less premature leaving from the former by the elimination of the potential early leavers from them.

(iv) *Four-Year Courses to G.C.E. at Ordinary Level in Grammar Schools.*

Four-year courses in grammar schools leading to, formerly the School Certificate, and now the G.C.E. examination at ordinary level are sometimes stated to encourage premature leaving from grammar schools, both by making it possible to leave with some kind of examination qualification at fifteen years

¹Chi-squared=1.906 $P < 0.20$.

²BURT, SIR CYRIL "Selection for Secondary Schools," this JOURNAL (1950), Vol. XX, pp. 3-8.

(iii) "Late-transference" between grammar and secondary modern schools and prevention of premature leaving from the former.

TABLE 2

"LATE-TRANSFERENCE" BETWEEN GRAMMAR AND SECONDARY MODERN SCHOOLS AND PREVENTION OF PREMATURE LEAVING FROM THE FORMER.

	Groups I and II "Better" ¹			Groups III and IV "Poorer"		
	Number of Pupils ²		No. of L.E.As.	Number of Pupils		No. of L.E.As.
	To Gr.	from Gr.		To Gr.	from Gr.	
(1) None in either direction			3			2
(2) Into grammar: none out of	69		2	79		6
(3) Into grammar: very few out of	482	15	7	398	51	5
(4) Into grammar: 25—75 per cent. as many out of	185	68	7	98	32	5
(5) Into grammar: 75—200 per cent. as many out of	22	24	3	35	75	9
Totals	758	107	22	610	158	27
Average number of 13-year-old pupils transferred per L.E.A.	34.5	4.9		22.6	5.8	
Ratio of transfers per 100 13-year-old ³ grammar school pupils	5.3	.75		3.9	1.0	
Ratio of transfers to total 13-year-old pupils per thousand	13.3			7.9		

of age, and more subtly by the suggestion that completion of some kind is possible at this early age. Enquiries were therefore made of a sample of L.E.As. with either low or high percentages of premature leavers. Replies were received from thirty-eight and the results are analysed in Table 3.

¹Groups I and II have fewer P.L.s. (after allowance has been made for influence of percentages of entrants on early leavers) than have groups III and IV. The members of the groups used in this analysis are forty-nine of the sixty-three authorities belonging to the sample.

²Numbers of pupils transferred are based on information provided by L.E.As. for the years 1950 and 1951. The average of the two years has been taken; in some cases there was considerable variation in the numbers for the two years. "Negligible" or "very few" were stated in a few cases for transfers from grammar schools: in these cases an estimate of two has been used.

³Calculations are based on the information supplied in the Ministry Statistical Return No. 38 for 1950. The figures for numbers of 13-year-old pupils do not include fee paying pupils in such schools or direct grant, etc.

TABLE 3

FOUR-YEAR G.C.E. COURSES AND PREMATURE LEAVING

	Groups I and II	Groups III and IV
L.E.As. with 4-year Courses	7	6
L.E.As. without 4-year Courses	12	13

There does not seem any evidence from this limited investigation that 4-year courses to G.C.E. necessarily encourage premature leaving from grammar schools. This is not surprising, in that it is obvious from the replies to the enquiries that in most cases only a proportion of the pupils are involved and this proportion would presumably include the potential University candidates; the more unsuccessful and therefore likely early leavers would not be included.

This completes the review of various educational practices likely to influence the rate of premature leaving.

II.—GENERAL CONCLUSIONS

The results of these studies indicate that for some time to come at any rate premature leaving from grammar schools must be expected to go on at different rates in different parts of the country. Its alleviation would seem to depend more on the cultivation of appropriate attitudes towards education in the community¹ rather than on finer techniques of educational allocation. These attitudes will obviously be developed more easily in some areas than others and their growth may be hastened if some means is found to stress the importance of the obligation which parents undertake in accepting one type of education rather than another, and by building up the prestige of all schools. This seems likely to occur more rapidly when the schools are developed in clearly defined social units.

The suggestion that the entrance to grammar schools should be reduced to 5 per cent. of the school population,² while undoubtedly virtually eliminating premature leaving from some schools, would probably not do so from others. The schools of these "difficult" areas need special consideration and perhaps extra amenities. The strain imposed by increased social and educational mobility both within the school and in the communities which they serve, should be recognised. It may be argued that, in this generation, a fairly high degree of premature leaving is inevitable in areas where only a small proportion of the parents' generation enjoyed a grammar school education; the change in expectation from one generation with no experience of grammar schools to

¹ "Community" here includes parents, pupils, teachers and employers of labour. For evidence of the differences in attitude of working class and middle class boys in grammar schools see *Social Mobility in Britain* Chap. VI pp. 141-159, "Social Status and Secondary Education" by H. T. Himmelweit, and for evidence of the growth of awareness of the importance of parents' attitudes in "preventing wastage of pupils from secondary schools" one may note the unusual if not unprecedented occasion of the meeting of the Joint Under Secretary of State for Scotland with the parents of pupils at Dumfries, mentioned in the leading article of *The Times Educational Supplement*, Oct. 1st, 1954.

² ALEXANDER, W. P., "The Organisation of Secondary Education." *Education*, Vol. CIII, No. 2667, March 5th, 1954.

the second and third generation prepared to undertake a 4-years' and then a 7- or 8-years' course there, must necessarily be a slow one—although it is obvious that in some otherwise similar areas the pace goes much more quickly than in others. Any special techniques of administration or any special methods of the schools of the former might well be tried out in the latter.

III.—SUMMARY

The purpose of this study was to extend the investigation of the influence of certain conditions on premature leaving from grammar schools, to include the educational practices mentioned below. Most of the information relates to sixty-three L.E.As. only, but these form a fairly representative sample of the 127 English local education authorities, as regards social background and entrance and premature withdrawals from grammar schools.

(1) *Common methods of selection at 11+.*

Intelligence tests, standardised attainment tests, school records, essays, and interviews appeared indiscriminately in the selection procedures of L.E.As. with both high and low percentages of premature leavers (after due allowance had been made for the "influence" of percentages of entrants on leavers). It was impossible to demonstrate in this sample any statistically significant association between the incorporation of any one of these techniques and reduced premature leaving.

(2) *Provision of secondary technical education.*

The possibility of choosing between three instead of two forms of secondary education at 11 plus could not be shown to be associated with the elimination of misfits as exemplified by early leavers from grammar schools. On the contrary technical education at 11, under the L.E.As. of the sample, occurred slightly more frequently in those areas where premature leaving was frequent, (although the difference was not statistically significant).

(3) *Transference within the secondary educational system.*

"Late-transference" to grammar schools occurred much more frequently among the "better" L.E.As. than among the less successful ones while with transference from the grammar schools the reverse was the case although the difference was not so great. Transference to secondary modern schools did not appear in 1951 and 1952 to help much towards removing potential early leavers.

(4) *Four-year courses to G.C.E. at ordinary level.*

No evidence was found from a limited sample to support a popular conception that such courses tend to encourage premature leaving.

General Conclusions.

This study, and that reported in the previous article, provide evidence in support of the view that premature leaving is a phenomenon socially as well as educationally determined. Its decrease requires primarily the development of certain attitudes towards continued education, rather than the refinements of selection techniques and the adjustment of grammar school places. The

former may be necessary for other purposes and the latter certainly require alteration in some areas, the direction of change being dependent upon one's answer to two questions—whether it is better that some children be educated in a grammar school for only four years rather than none, or whether one favours a complete re-organisation of secondary education.

IV.—APPENDIX

Analyses of some of the data discussed in this article but using different criteria for the classification of the L.E.As.

(1) L.E.As. GROUPED ACCORDING TO PERCENTAGES OF PREMATURE LEAVERS ONLY.

	Group A (Lowest %)	Group B	Group C	Group D (Highest %)
Total number of L.E.As.....	32	32	31	32
No. of L.E.As. in Sample	18	15	12	18
No Standardised Attainment Test	4	4	2	3
No Intelligence Test.....	0	1	1	1
One Intelligence Test	6	7	2	6
Two or more Intelligence Tests ..	12	7	9	11
No School Reports or Assessments	2	1	5	1
No Essay	9	6	6	6
No Interview	11	4	8	6

None of these differences is statistically significant.

(2) L.E.As. GROUPED ACCORDING TO DEVIATION FROM REGRESSION OF PREMATURE LEAVERS ON ENTRANTS.

	Group (i) (Best)	Group (ii)	Group (iii)	Group (iv) (Poorest)
Total number of L.E.As.	32	38	36	21
Number of L.E.As. in Sample ..	19	16	17	11
No Standardised Attainment Tests	5	2	2	4
No Intelligence Test	0	0	2	1
One Intelligence Test	7	7	3	4
Two or more Intelligence Tests ..	12	9	12	6
No School Report or Assessment	3	1	3	2
No Essay	8	10	5	4
No Interview	9	7	9	4
No Secondary Technical Educa- tion	6	5	6	2
Provision of Technical Education at 13+	9	7	6	2
Provision of Technical Education at 11+	4	4	5	7
	Groups (i) and (ii) combined (Better)		Groups (iii) and (iv) combined (Poorer)	
Number of L.E.As. for which information on transference was available	26		23	
No transference	4		1	
Transference to but not from Grammar Schools ..	3		5	
Transference to and less than 25 per cent. as many from	11		1	
Transference to and from 25 per cent.—75 per cent. as many from	5		7	
Transference to and from 75 per cent.—200 per cent. as many from Grammar schools	3		9*	
	Groups (i) and (ii) combined		Groups (iii) and (iv) combined	
With Four-year courses in Grammar schools	5 out of 16		8 out of 22	

* The only difference approaching significance in this analysis, (2), is that between those L.E.As. with late transference but with none or very few going from the grammar schools compared with those having 25 per cent. or over transferred from them.

$\chi^2=4.485$ $P=\text{approx. } 0.03$.

CHILDREN'S CHOICES IN INDIVIDUAL ACTIVITIES IN THE JUNIOR SCHOOL

By E. W. HUGHES

Organiser of Research, Institute of Education, Leicester.

I.—*Introduction.* II.—*Aims.* III.—*Recording procedure.* IV.—*Classification and analysis.* V.—*Results of analysis of the record sheets.* VI.—*Questionnaires.* VII.—*Summary.*

I.—INTRODUCTION.

THE allocation of lesson periods in the junior school to free choice of activity, has been criticised from time to time, in the correspondence columns of newspapers and in informal discussion, on the grounds that the pupils waste valuable school time in casual play and game forms of activity and that such activity periods are the resort of the lazy teacher. On the other hand the supporters of the principle of free choice argue that such methods of work make considerable demands on the teacher, but that essentially they provide opportunities for the practice of conventional skills at a time suited to the child and consequently at a time when there is likely to be considerable response.

This present article is based on a report on a study of the choice of activity during Individual Activity Periods in a city junior school (School A) over three consecutive terms, supplemented by evidence from a county junior school (School B) over twelve weeks.

The work originally developed out of an approach to the Institute of Education at Leicester by Miss E. E. Thornton, the Head Mistress of School A, and received support later from the Head Master of School B, Mr. I. Evans, in the form of a parallel research, which was conducted before the results from School A were known.

Unless stated to the contrary, the article will refer to School A. This is a school of some 500 boys and girls, situated in a good working-class district of the city, but bordering on a "highly desirable" residential area from which it draws about 50 per cent. of its pupils. From the intellectual point of view, the school could be classed as good. The accommodation and staffing ratio are considered normal for the city. The school has easy access to playing fields, swimming baths, cinema, and public library, and is in an area well provided with churches and chapels offering Sunday School facilities, Cub and Brownie packs. Good shopping facilities exist in the area and the school is within 8-10 minutes' bus ride from the city centre.

The Head Mistress and members of the staff subscribe to the principle that the aim of the school is to help the individual child to develop its capacities to the full. Accordingly, there is as much individual work and freedom as is consistent with a reasonable training in the essential basic subjects and in acceptance of responsibility. The school reflects this attitude throughout and the implementation of its policies is such that the school is referred to as an "Activity School," by parents and other teachers outside, in contrast to other junior schools near the area and elsewhere.

Nevertheless, apart from the methods of working in Discovery Subjects* and School Clubs, only one period per week of approximately one hour is devoted to "free activity" in the IIIrd and IVth years, with a slightly higher proportion in Years I and II. It is the work of these individual activity periods which forms the basis of this present study.

*Discovery Subjects—finding out about Countries and People, Natural History and Science.

II.—AIMS.

The aims of the research were :

- (a) To examine whether the individual activities chosen fell into identifiable categories (conventional or otherwise).
- (b) To consider to what extent factors present, or deficient, in the school, or larger environments had influenced the choice of activity.

III.—RECORDING PROCEDURE.

The first task was to establish a method of recording the actual activities carried out over a period of time. It was realized that the members of the staff were already fully occupied in dealing with the demands of the different groups and individuals and consequently it was decided that, as far as possible, the records should be maintained by the pupils themselves.

Two kinds of Record Sheet were designed. Type (1) was a quarto sheet of paper ruled in rows and columns to enable the child to enter brief descriptions of activities down the first wide column and the dates at the heads of the columns, along the top row. Crosses were entered in the appropriate cells in the rows to show persistence of given activities and change of activity was shown by change from row to row.

Record Sheet (1) contained 5 rows and 11 date columns.

Type (2) was a quarto sheet divided into 12 large cells. One cell was allocated to each activity period and the date and activity or activities were recorded in the space provided.

Pilot trials in School A showed that Type (1) was generally acceptable and convenient for one term's work. Type (2) was used exclusively in School B.

Generally, the Work Sheet recorded work done and degree of persistence of activity. Since it is on the nature of these entries that the value of this report depends, it is well to bear in mind the following reservations :

- (1) Under the conditions of the research we could not record *all* the individual activities in the school, although we did our best to get full records over a considerable period of time.
- (2) Occasionally a child may not, in fact, have followed out the activity recorded. In this connection we recommended a system of initialling by the teacher of the individual work sheets, at least twice a term, at random, as a check. Again the child's entry is at least one which he expects would be reasonably acceptable to the teacher as representing what he might have followed out.
- (3) In practice we always limit the environment in which the child works, to some extent, and hence free choice is limited to what is possible under existing conditions—limited free choice.
- (4) The entries on the forms would involve writing and spelling problems, especially in Year I.

IV.—CLASSIFICATION OF RESULTS AND METHOD OF ANALYSIS.

It was found possible to interpret the entries for children from all years I to IV. In the first term, only one set of records out of ten was obviously teacher-recorded and from time to time an odd sheet would appear similarly completed. Generally speaking, there seems to be no great obstacle to self-entry records, especially after Year I.

In the pilot analysis work, it was soon obvious that the bulk of the activities recorded could be classified in the conventional categories—Reading, Writing, Drawing and so on. The small amount of space left for writing on the Record

Sheet occasionally rendered an entry ambiguous. For example, a story title did not indicate whether the person was reading that story or writing a story. A later three-group broader classification into Arts and Crafts, Play and Game Forms, and Literate Pursuits, helped to eliminate errors from some of these ambiguous entries.

Two forms of analysis were adopted. The first amounted to abstracting the descriptions of the actual activities written on the work sheets and recording them in appropriate columns, labelled Reading, Art, etc., on a large analysis sheet. A system of tallies was used to show frequencies when the same activity turned up again during the term for the same or different children. This was a classification stage. The second amounted to entering the initials of the children in a column on the analysis sheet and entering in rows for the different children a series of letters, e.g., R—Reading ; W—Writing, etc., showing the categories of their various activities, day by day. (See Table I).

TABLE I
SAMPLE OF MAIN ANALYSIS SHEET
(Showing classes of activity engaged in by individuals during one term.)

	Initials	Date																Individual Range of Activities
YEAR I (Boys)	B.A.	A	A	M	A	M	A	M	A	M	M	M	A	M	M			A—6 M—8
	P.B.	P	P	A	M	M	M	Ac	W	W	W	M	M	M	P			A—1 M—6 Ac—1 P—3 W—3
	M.B.	Ac	M	M	M	M	M	A	A	M	Ac	M	M	M				A—2 M—9 Ac—2
	R.C.	A Ac	Ac	M	M	M	A	M	A	W	M	Ac	A	M	M	Ac		A—4 Ac—4 M—7 W—1
	K.D.	M	M	M	M	M	M	M	A	M	M	M	M	M	M	M		A—1 M—14
	etc.																	
	Daily Range of Activities	A R M P W Ac	3 2 12 3 — 3	2 2 12 3 — 2	4 1 12 — 1 1	2 — 17 — 1 1	— 3 17 — — 1											Grand Total All Activities In Term

A—Art ; M—Making things ; P—Play and Game Forms ; R—Reading ; W—Writing ; Ac—Acting.

The first procedure enabled us to see the extent of the variation of activities within one category :

e.g., the different names of books or reading material indulged in.
It also showed us the extent to which others tended to follow out precisely the

same activity. It is, however, impersonal and shows no reference to time of occurrence.

The second type of analysis sheet (Table I), enabled us to show :

- (a) Individual persistence in certain classes or categories of activity, over periods of time—see totals of letters in rows (Note—a double line of entry shows frequency of change of activity by moving from one line to another).
- (b) The daily range of classes of activity in a group as a whole—see totals in the appropriate column for that day.
- (c) The total range of activity throughout one term or period—see grand total of rows or columns.

It cannot, however, show variation of activity within one category of activity.

Thus, both systems of analysis are necessary, since the first is more likely to show precise deficiencies within our provision of activity material, and the second is likely to show up individual differences and consequently individual sources of satisfaction.

V.—RESULTS FROM ANALYSIS OF RECORD SHEETS.

Systematic recording of individual activities was obtained over the three terms, Summer Term, 1952, Autumn Term, 1952, Spring Term, 1953. The total activities recorded was 6,804 (1914, 2,021, 2,869). Owing to the annual transfer of classes in the Autumn, this represents continuous recording for 350 children over three terms and part recording for the remainder. This total represents only part of the actual individual activities which were engaged in. The teacher of Class IIIA (Summer Term, 1952), estimated that 70 per cent. of the activities for the class were actually noted for that particular term (although this figure would be higher if absence was allowed for). However, many of the class records show regular recording intervals (i.e., dates in regular sequence) and, in particular, Class IY (Session 1952/53) made a special trial on teacher recording and pupil self-recording and produced very full records. In general, we can look upon the entries as a very full recording of the actual individual activities for the period of time concerned. The activities recorded in School B over the twelve weeks numbered 2,626.

Overall Picture.

The raw material was analysed and presented in many different ways, but for the convenience of this article, a number of abstracts have been made. The plan of analysis was to regard the results for the separate terms and the separate classes as material from repeated researches and to look all the time for evidence of common tendencies and repeated tendencies. In this account, the overall pictures for the three terms will be contrasted and the cumulative figures broken down until we focus attention on the daily range of activities within a single class and the extent of individual variations of activities over a term.

The distributions of all the activities recorded for the three separate terms are shown in Table II. Here the interesting features are the predominance of Art and Craft activities and the relatively slight proportion of Play and Game Forms. In an attempt to bring out this contrast the activities have been re-grouped in a three-fold classification, Play and Game Forms, Art and Craft, and Literate Pursuits. The result is most interesting (see Table III). It has been common to criticise free choice of activity on the grounds that it leads to waste of time, yet for three consecutive terms the percentage of Play and Game Forms was less than ten and this tendency was repeated for either sex and maintained in the various age groups in many of the analyses for the three terms. In School

B the cumulative figure was 17 per cent., but the proportion of Play and Game forms fell off rapidly in the older pupils. Thus, in both schools a very large proportion (over 80 per cent.) of the activities were conventionally 'respectable' Art and Craft Activities or Literate Pursuits.

TABLE II
DISTRIBUTION OF ACTIVITIES IN SUBJECT CATEGORIES FOR THREE TERMS IN SCHOOL A
WITH CORRESPONDING STATISTICS FOR SCHOOL B.

SCHOOL A		Art	Mak- ing things	Act- ing	Read- ing	Play and Game forms	Writ- ing	Arith- metic	Dis- cov- ery	Misc- ellan- eous	Totals
SUMMER TERM, 1952	All Activities..	600	509	219	217	138	126	55	42	8	1,914
	Boys	255	365	83	103	84	39	22	27	7	
	Girls	345	144	136	114	54	87	33	15	1	
AUTUMN TERM, 1952	All Activities ..	501	382	92	444	158	126	124	159	35	2,021
	Boys	175	191	70	203	113	59	67	141	6	
	Girls	326	191	22	241	45	67	57	18	29	
SPRING TERM, 1953	All Activities ..	699	735	194	475	208	142	160	230	26	2,869
	Boys	255	435	50	250	103	44	88	169	13	
	Girls	444	300	144	225	105	98	72	61	13	
School B											
1952/1953 (Twelve Activity Periods)	All Activities ..	934	619	106	333	413	77	44	60	40	2,626
	Boys	516	232	41	194	240	28	37	43	15	
	Girls	418	387	65	139	173	49	7	17	25	

TABLE III
TERMS CONTRASTED.

School A	Play and Game Forms	Art and Crafts	Literate* Pursuits	Total Activities
Summer Term, 1952	8%	58%	34%	1,914
Autumn Term, 1952	9%	44%	47%	2,021
Spring Term, 1953	8%	50%	42%	2,869
*Literate Pursuits—includes here 3Rs., Acting and Discovery.				
School B	17	60	24	2,626

(It is not suggested that these Play and Game Forms are necessarily inappropriate in the school situation any more than the 'respectable' pursuits are necessarily appropriate.)

Terms Contrasted for School A

In contrasting the results for the three terms the most noticeable feature is the swing away from Art and Craft pursuits towards Literate Pursuits during the second term. This change seems to be centred in an increase in Arithmetic and Discovery Activities and an extraordinary increase in Reading mainly at the expense of Acting and Art and Craft Activities (see Tables II and III). Unlike the second term, the third term shows a large increase in total activities recorded (2,021—2,869). This seems to have resulted in a strengthening of the Literate Pursuits all round and a pronounced numerical increase in Art and Craft activities, although the actual proportion is not as high as in the first term.

Age Groups.

In attempting to locate these changes in emphasis, it is necessary to break down the overall picture into age groups and classes. Although the second term was the beginning of a new session, there is no evidence in the detailed statistics to suggest that the increase in emphasis on Literate Pursuits was due to the new intake in the first year. It appears to be more in the nature of a changed attitude in the children in Years II and III, IV of the new session. The statistics show that the boom in Reading was very much influenced by Year II (girls more than boys) and the Arithmetic increase was centred in Year IV (boys and girls) temporarily. The increase in Discovery subjects was also located in Year IV (mainly boys) and was maintained in the third term when it became an outstanding feature of the 4th Year activities for boys, in contrast to Art Activities for the girls.

TABLE IV
ACTIVITIES ACCORDING TO AGE AND SEX (%)—School A

	Session 1951/2			Session 1952/3					
	Summer Term, 1952			Autumn Term, 1952			Spring Term, 1953		
	Play and Game Forms/Misc.	Art and Crafts	Literate Pursuits	Play and Game Forms/Misc.	Art and Crafts	Literate Pursuits	Play and Game Forms/Misc.	Art and Crafts	Literate Pursuits
Boys :				Yr. I 0					
Year I	11	71	19	II 18	81	19	1	68	30
Year II	11	57	33	III 27	34	47	12	39	49
Year III	7	49	43	IV 0	32	41	25	41	33
Year IV	3	64	33		30	70	0	22	78
Girls :				Yr. I 0					
Year I	9	50	39	II 7	74	27	6	66	29
Year II	1	68	31	III 11	40	53	10	24	68
Year III	8	41	51	IV 6	76	11	16	55	29
Year IV	2	57	42		44	50	1	54	45

The statistics have been presented to allow for the transfer in the Summer—350 pupils (approximately) common to all three terms.

Generally speaking, there was a marked concentration of Art and Craft Activities in the classes in Year I and this proportion fell off in the older age groups. There was a corresponding tendency for Literate Pursuits to increase with age, but certain classes seemed to maintain characteristic features which disturbed this tendency to some extent.

The statistics from School B showed the reverse tendency for the proportions of Art and Craft activities to increase with age for both boys and girls and for the first year to be spent in more general experimentation.

Sex Differences.

The figures in Table IV show that large sex differences did not emerge in general between the broad categories Play and Game Forms, Art and Craft, Literate Pursuits. Within the category Literate Pursuits, Table II shows notable differences in Discovery Subjects and Acting. The most persistent difference in the various groupings was for girls' Art activities to predominate over their Craft activities and *vice versa* for boys in School A. In School B, however, boys' Art activities were more than double their Craft activities.

The sharply contrasted results in the preceding sections show the need for repeated research both within one school and in different schools and the need to relate these activities to environmental pressures—pressure from teachers, parents, friends, conditions, proximity of selection examination and the new session with fresh enthusiasms, and so on. The latter part of this article discusses attempts to estimate the effects of some such environmental factors.

Daily Range of Activities within a Single Class and Individual Variation in Activities over a Term.

So far, we have concentrated on the general picture presented by the classification of entries into recognisable and conventional categories, Term by Term, Class by Class and according to sex. It is useful now to consider the variation of activity within a single meeting of a class and also the extent of the variation in choice of activity for the individual child during one term.

In a previous section (Section IV) of this article, it was shown that the tabular form of class analysis enabled the range of activities within a given "lesson period" to be noted by observing the totals at the foot of the columns and the individual child's range of activity classes could be studied by observing the totals in the rows. See Table I.

The figures for the different age groups have been collected from the various record sheets and Tables V and VI give the pooled results for the three terms for School A with related figures for School B.

The statistics in Table V show that the number of different classes of activity proceeding during any one "lesson period," could range from two (rarely just one) to six or seven. Now it must be emphasized that this means classes of activity, like Art, Reading, Discovery, and the like. If we consider the fact that each of these classes contains many different topics, e.g., Art—pencil drawing, painting, crayon, etc., it can be seen that the number of separate activities could be very large, but for ease in considering the situation, it is convenient to think of groups of activity, in which case the range is up to seven, with a central tendency in the region of four. This is true for both boys and girls and was repeated in the parallel research in School B.

TABLE V
NUMBER OF DIFFERENT CLASSES OF ACTIVITY PROCEEDING DURING ANY ONE "LESSON PERIOD"

School (A)	Number of different classes of activity								Number of different classes of activity						
	1	2	3	4	5	6	7		1	2	3	4	5	6	7
Boys :								Girls :							
Year I ..	—	2	13	20	12	2	—	Year I ..	1	4	12	16	6	5	—
Year II ..	—	14	25	21	12	5	—	Year II ..	—	5	14	24	9	6	1
Part III ..	—	1	10	13	11	7	1	Part III ..	2	3	12	13	12	2	1
Part IV ..	3	6	11	11	9	9	—	Part IV ..	—	8	15	15	8	3	—
TOTAL ..	3	23	59	65	44	23	1	TOTAL ..	3	20	53	68	35	16	2
School (B)															
TOTAL ..	—	5	24	25	11	8	—	TOTAL ..	—	3	17	26	20	6	1

This set of figures gives the reader some idea of the extent of the provision and general supervision which such individual choice of activity calls for from the teacher. It is a sobering thought for some of the critics of individual choice of activity who have regarded such work as easy for the teacher. In this connection, it is interesting to note the following list of material and apparatus which was available for use in Class IIIA (Autumn Term, 1952) :

Painting Equipment.
Normal Text Books.
Reading Books.
Use of Library.
Materials for book-binding.
Materials for modelling
(card, glue, etc.).
Chess Sets.
Draughts.
Cards.

Microscope and slides (also instructions for same).
Film-Strip Projector and blank slides for mounting own material.
Experimental material (guided work) for expansion, contraction, use of heat for sterilising. (Experiments built up after suggestions by two members of class.
Materials for other occupations, such as knitting and sewing brought by children.

Common Tendencies in Activity during one "Lesson Period."

A second examination of the totals at the foot of the columns of the main analysis sheets (Table I) shows that there was a distinct tendency (60 per cent.) for a "modal" or common activity to turn up over and over again. That is, the statistics show a bunching in some given class of activity, e.g., Art, which sometimes persists for several lesson periods and then is replaced by another common activity, e.g., Acting. Thus, in spite of the periods being assigned to free choice of individual activity, there was a strong tendency for the members of the group to engage in a common class of activity.

The comparable figure for the parallel research in School B was 57 per cent.

TABLE VI
RANGE OF ACTIVITY FOR THE INDIVIDUAL CHILD DURING THE TERM.

School A.																	
Number of classes of activity chosen per term.									Number of classes of activity chosen per term.								
	1	2	3	4	5	6	7			1	2	3	4	5	6	7	
Boys :									Girls :								
Year I ..	14	46	41	13	4	1	—		Year I ..	10	21	31	21	4	—	—	
Year II ..	21	36	36	18	9	1	1		Year II ..	7	35	39	33	14	4	—	
Year III ..	15	39	37	23	5	1	1		Year III ..	14	40	28	24	9	1	—	
Year IV ..	25	42	27	11	3	—	—		Year IV ..	21	49	33	10	6	—	—	
TOTAL ..	75	163	141	65	21	3	2		TOTAL ..	52	145	131	88	33	5	—	

School B.																	
	1	2	3	4	5	6	7	8		1	2	3	4	5	6	7	8
TOTAL ...	4	28	47	30	14	3	1	1	TOTAL ..	—	20	37	37	18	3	1	—

The totals at the end of the rows of Table I show the different classes of activity engaged in by the child during one term, and the frequency with which each class was repeated, e.g., B.A. tried two different classes of activity, while P.B. tried five different classes. The figures in the table above (Table VI) are the collected statistics for three terms to show this range of individual choice of class of activity. The range is up to seven classes of activity for boys and six for girls. There is a central tendency in the region of two to three classes for both sexes. This, of course, does not mean that there was not a big range of activities within one category (say Art) but it does mean that there was a certain persistence in choice of given classes of activity by individuals.

The parallel research in School B gave a somewhat similar result, see Table VI.

If repeated research under widely different conditions confirm these two tendencies (1) for a common class of activity to be a frequent feature of the daily activity periods, and (2) for individual children to tend to choose from a narrow range of classes of activity, then individual activity periods in the junior school become a valuable source of evidence of Educational Interests with corresponding significance in Education Guidance at 11+.

VI.—QUESTIONNAIRES.

In view of our pilot work in this field, it was decided to introduce some questionnaires into the research in order to try to estimate the extent of school and other environmental pressures on the choice of activity. At this stage of reporting, the questionnaires referred more to the human elements rather than the material elements, in the environment. Two questionnaires were used. The first asked for a direct statement about the interests of parents, teachers, friends (and self) in that these may have had some bearing on the child's present activity. The answers given would not necessarily be accepted by the persons concerned, but were what the child thought appropriate to record at the time.

The second questionnaire was designed to remedy some of the defects of the first and asked for child's activities pleasing to parents, teachers, friends, self. The technique was fairly simple. The children were asked :

" Which would please him/her most, that you should be able to—

write well
spell well
make-up stories

like writing, etc."

Parents, head teacher, teachers, friends, self, anyone else, were considered one at a time and the names, father, mother, etc., were written against the activity (ies) deemed to be pleasing to that person. Questionnaire (2) was believed to be more likely to provide evidence of a causal relationship with activities than Questionnaire (1).

These questionnaires were set to the 350 children for whom continuous recordings were available over three terms. The questionnaires were set in December, 1952, and are relevant to both the Summer terms and the Autumn terms in that some of the questions refer to both terms.

It might be wondered to what extent a young child of eight years or so would be capable of such analysis of the behaviour of others, but it is argued that the child is very much aware of the approval and example of adults and school fellows, or at least, so we tend to assume in life as a whole.

The teachers who gave the questionnaire and explained the instructions did not report any major difficulties and the completed scripts were easily decipherable. (Questions, particularly embarrassing to some children, e.g., parental relationship, were avoided by telling the children to put a large "X" if they could not answer the questions in Questionnaire (1).)

The results from the 700 questionnaires were pooled in some cases for general impressions and later separated on the bases of Questionnaire (1) or (2), Sex Difference, Age and Class Differences. In the interests of economy in this article, a selection of the analyses only will be presented. Occasionally, this will be misleading because in this part of the research in particular, it is shown that the result depends very much on the nature of the group under review—see Tables X and XI.

School Influences—Head and Staff.

Table VII shows the pooled results for members of the Staff and Head from both questionnaires. The entries, therefore, represent the answers to the questions :

What is your present teacher interested in ?

What was your teacher last term interested in ?

What activity of yours would please the teacher ? (present class teacher and any other).

and similar questions relating to the head mistress.

The method of analysis is likely to underestimate rather than overestimate the three categories of activity, Discovery, the three Rs., Art and Crafts, as allowances were made when combining separate entries from Questionnaire (2) under these multiple headings.

Generally, the results show that the children's estimates of teacher interests are well spread over the various categories of activity with a noticeable emphasis in favour of the 3 Rs. and only moderate pressure in favour of Music. Further analysis into profiles for individual teachers show that the children see these

teachers as quite different individuals, some appearing in their specialist capacity only, e.g., the pressure in favour of music seems to be centred chiefly in one man.

The head mistress, perhaps traditionally characteristically for school heads, stands out in favour of the 3 Rs. and the School Society with a strong supporting background which renders the whole reminiscent of the Staff pooled results.

TABLE VII
SCHOOL INFLUENCES IN FAVOUR OF SPECIFIED ACTIVITIES—AS JUDGED BY
YEARS II, III AND IV (CHRISTMAS, 1952)

	No.	Sex	Three R's	Games and P.T.	Dis- covery	Art and Craft	Social School Soc.	Music	Un- known Don't know	
Members of Staff as seen by Years II, III, IV Christmas, 1952	1	F	32	25	12	31	11	—	8	
	2	M	21	51	16	33	12	—	7	
	3	F	30	3	8	11	13	1	3	
	4	M	20	15	35	17	15	3	4	
	5	F	54	34	77	27	35	4	11	
	6	M	42	18	51	12	17	4	26	
	7	F	16	15	5	35	8	5	2	
	8	M	67	47	56	26	18	16	24	
	9	F	38	77	19	28	51	10	16	
	10	M	62	27	21	7	6	1	8	
	11	M	—	1	—	—	—	89	—	} Not reported on as "class teachers"
	12	F	41	13	15	10	8	10	2	
	13	M	13	11	6	29	5	8	5	
	14	M	9	—	1	13	—	—	—	
	15	M	19	5	7	9	7	—	4	
All members of Staff			464	342	329	288	206	151	120	
Head Mistress			311	104	133	74	265	85	101	

In interpreting these results, it should be remembered that one section of the school would be taking the selection examination for Secondary Education early in the following term and the questionnaire results and activity results may reflect a general sense of urgency in matters of the 3 Rs.

Home Influences.

Three sets of results are available in this section :

- Children's ideas of Parents' Own Interests.
- Children's ideas of Parents' Work Activities.
- Children's ideas of their own Activities which would please the parents.

Each of these analyses stands in its own right. We are suggesting that these represent factors present in the child's home environment which *may* have an effect on the trend of activities reported in the earlier sections. In thinking of casual relations we feel that the positive lead comes from Questionnaire (2), "Which of your activities would please your Parents?" In the other two cases, Parents' Own Interests and Parents' Jobs, we can only show degree of association.

Parents' Own Interests (as judged by children).

The first feature of note is that both parents are seen by both boys and girls to be interested in their own (parents) Work Activity. Another peak appears in Art and Craft Pursuits, and this is unusually predominant since it includes

knitting, sewing, woodwork and painting, some of which could easily be Work rather than Art/Craft activities. Father is reported as more interested in Literate Pursuits than Mother and both boys and girls see the Father's interest in Discovery Subjects. In view of the number of "X" entries, the present statistics, in more detailed analyses, suggest that the girls are more certain of their Parents' interest preferences than the boys. (Table VIII).

TABLE VIII
CHILDREN'S IDEAS OF PARENTS' OWN INTERESTS.
(Questionnaire 1)

Category of Activity	Father's Interests			Mother's Interests			Both Parents
	Boys	Girls	Boys and Girls	Boys	Girls	Boys and Girls	Boys and Girls
Own Work Activity	36	30	66	51	50	101	167
Art and Craft	31	27	58	32	75	107	165
Literate Pursuits	24	40	64	25	21	46	110
Discovery and Gardening	29	22	51	13	2	15	66
Social, including T.V. and Radio	13	14	27	8	9	17	44
Games and Physical Activities	23	14	37	4	2	6	43
Music and Speech	4	9	13	6	6	12	25
Possessions	5	5	10	2	—	2	12
"X" or "?"	35	19	54	50	16	66	120

Parents' Jobs (as reported by the children).

The previous section suggested that many parents had a special interest in their own work activities—what are these work activities?

In 80 per cent. of the cases it was possible to classify father's job in broad occupational categories. The greatest single entry was in Industry, as would be expected, but this figure together with Agriculture, ranks at about 40 per cent. (It also includes some supervisory, management and technical personnel.)

The remaining 40 per cent. are in "white collar" jobs. This is an unusually large proportion of fathers in occupations which can be expected to call for "more than average" literate ability. The family discussions round such work cannot fail to convey the importance, or essential nature, of literate ability in the work situation and thus constitute one more pressure in favour of the 3 Rs.

TABLE IX
CHILDREN'S IDEAS OF PARENTS' WORK ACTIVITIES
(Combined Results.)

Job Category	Father's Job	Mother's Job
Industry/Building	130 } Approx.	26 } Approx.
Agriculture	15 } 40%	0 } 22%
Commerce, including Distribution	51	16 } known
Professions	30	18 } "out to
Services (Local, Armed, Social)	41	11 } work."
Office Work	11	5
Don't Know	69—20%	73—21%
Home Duties/Domestic	1	200—57%
TOTAL	348	349

Only 22 per cent. of the mothers are definitely stated as "Out to Work." This is low for this particular town. Once again, at least two-thirds of these are in non-industrial occupations. Some 57 per cent. of the mothers are engaged in Home Duties. This latter is likely to have a neutral (expected) effect, so that the influence from Mother's job is again likely to come from the 22 per cent. and be in favour of the 3 Rs. (There are quite a number of teachers on both sides. Table IX.)

TABLE X
CHILDREN'S IDEAS OF THEIR OWN ACTIVITIES PLEASING TO PARENTS.
(Pooled results from Questionnaire (2).)

	Social Behav- iour	3 R's	Games and P.T.-	Disc. Subj.	Art	Craft	Music
Both Parents as seen by Boys and Girls :	335	317	234	228	202	180	177
Year II only	107	123	81	81	77	51	74
Year III only	94	96	74	81	65	59	58
Year IV only	134	98	79	66	60	70	45
Both Parents as seen by Boys only	162	158	121	115	96	86	76
Both Parents, as seen by Girls only	173	159	113	113	106	94	101

TABLE XI
ACTIVITIES PLEASING TO PARENTS
(Detailed results from Questionnaire (2).)

	Activities pleasing to Father							Activities pleasing to Mother						
	Social Behaviour	Three R's	Games P.T.	Discovery Subjects	Art	Craft	Music	Social Behaviour	Three R's	Games and P.T.	Discovery Subjects	Art	Craft	Music
As seen by Boys :														
Year II	14	23	16	18	17	14	10	31	35	17	19	19	11	17
Year III	16	21	20	24	11	19	10	36	30	21	18	13	11	24
Year IV	22	20	27	19	18	20	5	43	29	20	17	18	11	10
TOTAL (Boys)	52	64	63	61	46	53	25	110	94	58	54	50	33	51
As seen by Girls :														
Year II	26	32	22	23	20	12	19	36	33	26	21	21	14	28
Year III	14	18	15	19	21	15	10	28	27	18	20	20	14	14
Year IV	27	22	15	19	12	11	14	42	27	17	11	12	28	16
TOTAL (Girls)	67	72	52	61	53	38	43	106	87	61	52	53	56	58
GRAND TOTAL— BOYS AND GIRLS	119	136	115	122	99	91	68	216	181	119	106	103	89	109

Children's Ideas of their own Activities Pleasing to Parents.

The figures in Table X (pooled results) suggest that the children thought that parents were strongly in favour of the 3 Rs. and good Social behaviour, with some secondary preferences in Games and Discovery Activities. The same picture is supported by the statistics for Years II, III and IV, separately.

A re-analysis into the views of boys and girls, separately, gives support to this growing conviction.

If, however, we now show boys' views of father and mother separately, and girls' views of father and mother separately (see Table XI, detailed results), our confidence in activities pleasing to mother is re-inforced—both boys and girls select 3 Rs. and Social behaviour. Father, however, assumes a much more dispersed and diverse set of pleasing activities, especially as seen by boys.

If we now continue the analysis to allow for various age groups and classes, we find that father's preferences are not seen consistently throughout the various classes, but that mother appears most consistently in favour of the 3 Rs. and Social behaviour.

The boys in particular see this characteristic in mother, but father they view as much more diverse.

Questionnaires were not set in the parallel research in School B, but the Head listed his impressions of the emphasis which he and his staff could be said to give. There was a strong element in favour of Art and Craft activities. Again, the Head produced an estimate of the occupation analysis of the parents. There was an extremely high proportion of parents in manufacturing industry as craftsmen and tradesmen. The academic level of the school was estimated at average for the county. In view of these observations, it is interesting to note that the feature of School B was narrowing of the range of activities from a diverse array in the early classes to Art and Crafts pursuits in the older age groups.

VII.—SUMMARY.

This article is concerned with a study of individual activities under conditions of free choice of activity in a City Junior School (A) and a County Junior School (B). 6,804 individual activities and approximately 700 questionnaires were analysed from School A and 2,626 individual activities from School B. The results show that :

- 1.—It was possible to classify the bulk of the activities in conventional categories—Reading, Writing, Art and the like.
- 2.—The predominant group of activities was often Art and Craft for both sexes. This was particularly true of the younger age groups in School A, the Art and Craft activities giving way to Literate Pursuits in the older age groups. School B showed a reverse tendency for Art and Craft to increase in the older age-groups.
- 3.—In School A, Play and Game Forms of Activity formed less than 10 per cent. of the total activities for the term and this was repeated over three consecutive terms. In School B, the corresponding figure was 17 per cent., with a tendency for Play and Game Forms to fall off in the older age groups. Thus, over 80 per cent. of the activities recorded were Literate Pursuits or Art and Craft forms.
- 4.—During any one "lesson period," there was a range of up to seven different classes of activity (e.g., Reading, Art, etc., with all their variations) going on at times and a central tendency in this respect of four for both boys and girls and in both schools. This reminds us of the large demands which

individual activity periods make on the Staff by way of provision and supervision.

5.—There was evidence of a tendency for a noticeable proportion of a group to engage in a common class of activity in about 60 per cent. of the "lesson periods" studied and there was a suggestion of individual persistence in activity, especially in the older children, to the extent that they chose activities from two to three classes of activity only. This has implications for educational guidance at 11+.

The questionnaires, which were intended to estimate some of the school and local environmental pressures on the children in School A indicated considerable pressures in favour of the 3 Rs.

In particular, the questionnaires showed that :

1.—Whilst individual members of the staff were all viewed differently, they combined together to present a strong front in support of a wide range of activity with a hint in favour of the 3 Rs. The Head Mistress was reported upon, perhaps characteristically, for school heads, as in favour of the 3 Rs. and good social behaviour against a full supporting background of other interests.

2.—Parents "suggested" the importance of Literacy in their own activities, in their work activities and conversations about their work, and in the impressions which they conveyed of pleasing activities in their children. Mother was viewed most consistently in this respect throughout the various age ranges by both boys and girls.

ACKNOWLEDGMENTS.

The writer wishes to express his appreciation of the work which has been done by Miss Thornton and Mr. Evans and the Staffs of the two schools concerned in the collecting of material for this investigation. He is very grateful for the continued support over a long period of time and the active discussions on the planning and the difficulties arising.

ENGLISH COMPOSITION IN SECONDARY SCHOOL SELECTION

By J. D. NISBET

Lecturer in Education, University of Aberdeen.

I.—Introduction. II.—Sample and Measures. III.—Results. IV.—Summary and Conclusions.

I.—INTRODUCTION.

Previous investigations of the marking of English composition have been mostly concerned with the reliability of marks or with the consistency between examiners' assessments. Notable among such investigations is the following: "The Efficiency of Different Methods of Marking English Composition": CAST, B. M. D. (*Brit. J. Educ. Psychol.* IX and X).

A further question is important if English composition is used in selection for secondary education, the question of validity. Any measure used in secondary school selection should have not only high reliability but also high validity—that is, it should predict accurately a pupil's subsequent success in secondary school. The comparison of different methods of marking compositions should therefore include consideration of the validity of the methods, measured against the validity of the objective test of English attainment. As there have been few such validation studies of composition marks, some interest may attach to the results of an investigation¹ in which scores in objective tests and marks obtained by two methods of marking compositions were compared with estimates of success in secondary school three years later.

II.—SAMPLE AND MEASURES.

The sample consisted of all those born in Aberdeen in the month of April, 1939. There were 206 of these, but after losses (mostly due to migration) complete data were available for 169 cases. In November, 1950, the Education Authority's battery of selection tests was applied to this group; and as a result, in 1951, forty-one were promoted to senior secondary schools (the approximate Scottish equivalent of the secondary grammar school) and 128 to junior secondary schools. The selection battery consisted of Moray House Tests 41 and 42, Moray House English Test 19 and Arithmetic Test 19, and scaled teachers' estimates of English and Arithmetic attainment.

The compositions were written in February, 1951, with a time limit of 45 minutes and the following choice of topics:

- 1.—A day in the life of a postman or a policeman or a shopgirl.
- 2.—Suppose you have lost your pet dog or cat. Tell your adventures during the search for it.
- 3.—A day spent by the sea or among the hills.

¹ Several teachers collaborated in this inquiry, in particular Mr. J. Zamory, M.A., who supervised correction, and Miss A. F. Laing, M.A., Ed.B., who obtained criterion assessments. Their assistance, together with the facilities provided by Aberdeen Education Authority, is gratefully acknowledged.

Marking these compositions was done by two separate methods. Firstly, four markers used a "quick impression" method, handling approximately fifty compositions per hour, following the method described by Wiseman¹: the addition of their four marks gave a final mark for each paper. Secondly, a single marker assessed each composition subjectively, using an analytic scheme described by Schonell,² allotting a mark for thought or content and vocabulary, a mark for structure and a mark for mechanical accuracy: the addition of these three assessments gave a final mark.

The criterion used for success in the secondary school was a scholastic one, expressed in terms of performance in external examinations. This was necessary because the children were scattered throughout the city in different schools, some of which were senior secondary schools and some junior secondary. The teachers were asked to estimate the probable performance of the pupils in one or other of the two leaving examinations common to all the schools in the area, the Scottish Leaving Certificate and the Junior Secondary Certificate. Pupils in senior secondary schools were assessed on a three-point scale:

- (A) Seems certain to obtain the Scottish Leaving Certificate in the fifth year.
- (B) Has reasonably good chance of gaining the Scottish Leaving Certificate in the sixth year.
- (C) Is unlikely to gain the Scottish Leaving Certificate.

Pupils in junior secondary schools were assessed on a four-point scale:

- (a) Shows such ability that, if given the chance, would possibly have gained the Scottish Leaving Certificate.
- (b) Seems certain to gain the Junior Secondary Certificate.
- (c) Is likely to gain the Junior Secondary Certificate.
- (d) Is unlikely to gain the Junior Secondary Certificate.

Because of the coarse grouping of these separate scales, the two sets of assessments were combined arbitrarily into a single five-point scale to include the whole group, making (a) on the junior secondary scale equivalent to (B) on the senior secondary scale, and (b) on the junior secondary scale equivalent to (C) on the senior secondary scale. On this composite five-point scale, it will be seen that it was impossible for a junior secondary pupil to gain the highest assessment, and impossible for a senior secondary pupil to fall into either of the two lowest assessments. However, as is shown by the results below, the correlations between this composite criterion and the various tests were closely comparable with the correlations obtained from the senior secondary and junior secondary groups separately, when these were corrected for homogeneity.

III.—RESULTS.

The correlations between the composition marks and test scores and the criterion of secondary school success are given in Table I, first for senior secondary pupils only (41 cases), then for junior secondary pupils only (128 cases), and finally for the total group using the composite criterion. The coefficients for the senior and junior secondary groups were corrected for

¹WISEMAN, S.: "The Marking of English Composition in Grammar School Selection," *Brit. J. Educ. Psychol.* XIX, 200-209, 1949.

²SCHONELL, F. J.: *Backwardness in the Basic Subjects* (Oliver and Boyd, Edinburgh, 1942). Schedule O, p. 484.

homogeneity by the Otis-Kelley formula : the corrected coefficients are given in brackets.

TABLE I
CORRELATIONS BETWEEN TESTS AND SECONDARY SCHOOL SUCCESS.

Correlation between criterion of secondary school success and :	Senior Secondary group	Junior Secondary group	Total group
Composition :			
Quick impression, 4 markers35(.75)	.59(.69)	.73
Analytic method47(.78)	.60(.70)	.74
Test Battery :			
Moray House Test 4148(.79)	.56(.68)	.73
Moray House Test 4234(.78)	.60(.71)	.74
Moray House English Test 19.....	.52(.82)	.71(.83)	.84
Moray House Arithmetic Test 1954(.75)	.68(.78)	.79
Scaled teachers' estimates37(.73)	.60(.70)	.73

Since this validation study is based on a small number of cases, the differences between the correlation coefficients in Table I are not significant. It would therefore be unwise to draw definite conclusions from the relative size of the various coefficients, beyond stating that the results give general support to the validity of both methods of composition marking for selection purposes. However, in this sample at least, the objective test of English attainment proved superior to both sets of composition marks.

Further results may be given briefly. The reliability of the two methods of marking used was calculated by a re-marking of forty-one scripts (selected by alphabetical order of surname) after an interval of one month. The reliability coefficients were : for the quick-impression method—Marker A, .88 ; Marker B, .88 ; Marker C, .72 ; Marker D, .87 ; for the analytic method—.86. When the marks of the four quick-impression markers were combined the reliability coefficient was .96.

The validity of the marks of the separate quick-impression markers (i.e. the correlation between these marks and secondary school success) was : Marker A, .61 ; Marker B, .62 ; Marker C, .66 ; Marker D, .67. The quick-impression marks proved in fact to be more influenced by the number of words written than the analytic method : the correlation between marks and number of words was .51 for the combined quick-impression marks and .32 for the analytic marks. In the analytic method, the correlation between secondary school success and the marks allotted for content was .66 ; between success and structure, .66 ; between success and mechanical accuracy, .60. The correlation between the analytic marks and the combined quick-impression marks was .86.

Correlations with the objective test of English attainment (M.H.E.19) were : for the combined quick-impression marks, .82 ; for the analytic marks, .85. These coefficients represent a close agreement between composition mark and test score. The composition marks consequently are largely duplicating what is already being measured by the test battery. The multiple correlation between the test battery and secondary school success is .86 ; the addition of composition to the battery merely raises the multiple correlation to .87.

IV.—SUMMARY AND CONCLUSIONS

1.—Compositions written by a random sample of 169 children at age $11\frac{1}{2}$ were marked by two methods, a quick-impression method using four markers, and an analytic method using a single marker.

2.—The marks obtained by these two methods were compared with a scholastic criterion of success in the secondary school at age $14\frac{1}{2}$, based on teachers' estimates of the pupils' probable performance in common external examinations. Scores in objective tests, including an English attainment test, were also compared with this criterion.

3.—The validity of the composition marks, measured by their efficiency in predicting secondary school success, was satisfactorily high—comparable with several of the objective tests commonly used in the selection procedure. The number of cases was not sufficient to identify either method as superior.

4.—The English attainment test gave a higher (but not significantly higher) correlation with secondary school success than did the composition marks.

5.—In this sample there was no statistical evidence that the inclusion of composition in a selection test battery would markedly improve the predictive efficiency of the battery. It is recognised, however, that there are also relevant educational considerations which lie outside the scope of this inquiry.

6.—In addition there must be considered, of course, the extra cost of having four examiners for the quick-impression method.

EX-NATIONAL SERVICE AND "SCHOOLBOY" UNDERGRADUATES: A COMPARATIVE STUDY OF ACADEMIC PERFORMANCE.

By ELISABETH M. BROWN

Applied Psychology Unit, Department of Psychology, University of Edinburgh

AND

JAMES G. M'COMISKY

Applied Psychology Unit, Department of Psychology, University of Edinburgh

- I.—Introduction. II.—Study Design. III.—Findings. IV.—Discussion.
V.—Conclusions. VI.—Summary.

I.—INTRODUCTION.

Quite definite views are often met with among teachers and others regarding the advisability of young men doing National Service before going up to University. Nothing, however, as far as the present writers are aware, has been published which suggests that a systematic inquiry into this very important question has been made anywhere in the United Kingdom. The present study, which was originally undertaken on behalf of the King George Jubilee Trust Inquiry into Influences affecting Young People, is felt to have some value as a first step in meeting this need. It is, in effect, a follow-up of two groups of Ex-National Servicemen, one group in each of the Faculties of Science and Arts, who entered the University of Edinburgh in 1949 and in 1950, and a comparison of their academic performance with that of two equivalent groups of students who entered the University straight from school.

Details of the design of the Study and findings and conclusions follow.

II.—STUDY DESIGN.

1.—Seventy-three Ex-National Service Arts students from the 1949 and 1950 Entries were matched with seventy-three "schoolboy" students drawn from the Arts Entries of the same years. Similarly, seventy-four Ex-National Service students drawn from the 1949 and 1950 Entries to the Science Faculty were matched with seventy-four "schoolboy" students from the same Entries. Matching was done on the following points:

	Arts Faculty	Science Faculty
(a) Scottish family background,		
(b) Scottish Secondary education,		
(c) Scottish S.L.C. with same number of Highers		
and Lowers approximately, and		
(d) Same intended Degree Course at the University		
	TOTALS	58 66
(a) English Secondary education and		
(b) English School Certificate with approximately		
same level of attainment		
	TOTALS	15 8
	73	74
	—	—

Age-bands.—The age-bands from which the equivalent groups in the two Faculties were drawn were :

Schoolboys 17 years to 19 years 10 months.
Ex-National Servicemen 19 years 3 months to 23 years.

2.—Comparisons have been made between the equivalent groups in respect of : class of Degree awarded on graduation ; numbers abandoning Degree courses ; examination performance in subjects where the numbers taking the examination and the method of marking papers were such as to enable the comparison to be made effectively. Findings are tabulated in Tables I and II and conclusions indicated.

III.—FINDINGS.

TABLE I
RESULTS OF DEGREE COURSES

	EQUIVALENT GROUPS (ARTS)				EQUIVALENT GROUPS (SCIENCE)			
	Scottish and English		Scottish only		Scottish and English		Scottish only	
	Boys	Ex- N.S.	Boys	Ex- N.S.	Boys	Ex- N.S.	Boys	Ex- N.S.
<i>Hons :</i>								
1st Class	7	3	5	1	8	5	7	5
2nd Class	15	22	12	17	14	14	12	10
3rd Class	1	2	1	1	—	2	—	2
Unclassed	—	1	—	1	—	—	—	—
	(23)	(28)	(18)	(20)	(22)	(21)	(19)	(17)
<i>Ord. :</i>								
Full Degree	32	16	25	12	27	21	25	19
War Degree	—	5	—	5	—	2	—	2
	(32)	(21)	(25)	(17)	(27)	(23)	(25)	(21)
<i>Students still studying :</i>								
Honours	3	3	3	2	2	—	2	—
Ordinary	2	2	2	2	8	6	7	5
	(5)	(5)	(5)	(4)	(10)	(6)	(9)	(5)
<i>Abandoned Course</i>	13	19	10	17	15	24	13	23
TOTAL	73	73	58	58	74	74	66	66

NOTE.—There are no statistically significant differences between academic performance of Boys and Ex-National Service Students in either the Arts or Science groups.

Within the limitation that differences in over-all performance between the groups are *not* statistically significant, the results given in Table I indicate that :

Arts :

- (a) Ex-National Servicemen of the same level of Leaving Certificate attainment as Schoolboys going straight to University, tend to do better than the Schoolboys in terms of Class of Degree awarded : 38 per cent. of the Ex-National Servicemen obtained Honours Degrees against 31 per cent. of the Schoolboys.

- (b) Schoolboys, however, tend to do better than the Ex-National Servicemen in respect of:
- (i) Percentage of the remainder obtaining Ordinary Degrees: 44 per cent. of the total group in the case of the Schoolboys compared with 29 per cent. of the Ex-National Servicemen.
 - (ii) Number of students abandoning Degree courses: 18 per cent. of the Schoolboys as against 26 per cent. of the Ex-National Servicemen.

Science:

Schoolboys going straight to University and Ex-National Service students with the same level of Leaving Certificate tend to gain the *same* percentage of Honours Degrees: 29 per cent. of the equivalent group in the case of the Ex-National Servicemen and 30 per cent. in the case of the Schoolboys. Schoolboys, however, tend to do better than Ex-National Servicemen in respect of:

- (i) Percentage of the remainder obtaining Ordinary Degrees: 37 per cent. of the Schoolboys against 30 per cent. of the Ex-National Servicemen.
- (ii) Number of students abandoning Degree Courses: 20 per cent. of the Schoolboys against 34 per cent. of the Ex-National Servicemen.

Comparison of the figures for the years 1949 and 1950, taken separately, shows that there is considerable difference in consistency of relative academic performance by the groups as between the two Faculties. In Science the pattern of performance for the two years is broadly speaking the same, with the Ex-National Service group of 1950 entrants doing a little better than that of the 1949 entrants. In the Arts Faculty, however, the pattern as between the two years is quite different, the main change being due to the fact that a much higher percentage of the 1950 Schoolboys (43 per cent. for 1950 as against 23 per cent. for 1949) gained Honours, while the percentage of the 1950 Ex-National Servicemen gaining Honours remained substantially the same (37 per cent. of the 1950 group as against 39 per cent. of the 1949 group.)

TABLE II
EXAMINATION PERFORMANCE ON DEGREE COURSES
Marks are given as percentages.

	Boys			Ex-N.S. Men		
	Mean	S.D.	No.	Mean	S.D.	No.
<i>Arts Faculty:</i>						
English (1)	57.0	8.6	24	53.8	5.7	28
Latin (1)*	58.9	8.7	21	52.7	6.4	17
Maths. (1)	53.3	11.9	12	43.0	14.1	14
British History	52.2	7.1	48	52.9	7.5	44
Moral Philosophy	52.4	6.9	37	54.3	4.4	30
Political Economy ..	55.7	7.7	30	55.6	4.7	28
<i>Science Faculty:</i>						
Maths. (1)†	55.7	16.1	48	47.2	16.4	44
Maths (2)†	59.8	11.5	30	45.8	15.8	19
Chemistry (1)	59.0	11.6	66	58.1	9.1	64
Geology (1)	60.4	8.1	27	58.0	7.7	22

FIVE Schoolboys and SIX Ex-National Servicemen failed all 1st year courses in Arts.
SEVEN Schoolboys and ELEVEN Ex-National Servicemen failed all 1st year courses in Science.

* diff. sig. at .05 level. † diff. sig. at .01 level.

These results indicate that :

- (a) Ex-National Servicemen of the same level of Leaving Certificate as Schoolboy students do as well as the latter in some subjects which they have done at Secondary School, e.g., English and History ; they also do equally well in some " University subjects," e.g., Moral Philosophy and Political Economy.
- (b) Ex-National Servicemen's examination performance in Mathematics in both Faculties, and Latin in the Arts Faculty, is poorer than that of Schoolboy students who have the same level of Leaving Certificate attainment.

IV.—DISCUSSION.

A basic limiting factor in the inquiry must be emphasised. This is the tendency, particularly in Arts, for Schoolboys with better-class Leaving Certificates to go straight from school to University, usually on account of Scholarship award, and for those who do less well in their Leaving Certificate examination to go on National Service first. In effect, this means that the Ex-National Servicemen with their comparatively mediocre University Entrance qualifications can only be directly compared with Schoolboys with similar mediocre qualifications, rather than with " bright " Schoolboys or Schoolboy students in general.

Turning to the results it will be seen that, on balance, what differences there are between the two types of entrant are in favour of the Schoolboys. It will also be seen that the relative performance of the Ex-National Servicemen differs as between the two Faculties in that while in Arts they score a point over the Schoolboys by gaining a higher percentage of Honours, they do not score a single major point against the Schoolboys in Science. The Ex-National Servicemen in Science, generally speaking, lose parity with the Schoolboys through their lack of success in achieving Honours on the same relative scale as in Arts and because of their greater tendency, compared with the Schoolboys, to abandon their Degree Course completely following failure in examinations. The more stringent conditions of award of Further Education and Training Grants to Ex-National Servicemen compared with those of some of the Bursary awards to Schoolboys seem to be a factor in this latter connection.

With regard to examination performance as shown in Table II, statistically significant differences were found in favour of the Schoolboys as opposed to the Ex-National Servicemen in Mathematics 1st Course and 2nd Course in the Science Faculty, and in Latin 1st Course in the Arts Faculty. This comparative weakness of the Science Ex-National Servicemen in Mathematics seems important, since examination of the figures shows that of the twenty-four Ex-National Servicemen in the equivalent group who failed Mathematics 1st Course sixteen abandoned their Degree Courses completely and five fall in the " still studying " category. The comparative figures for Science Schoolboys in the equivalent group are thirteen failures in Mathematics 1st Course, eight of whom abandoned their Degree Courses completely with two " still studying."

There is the additional important point that in many Science Degree Courses, as compared with Arts, Mathematics is not only a compulsory subject but enters directly also into other basic Science subjects such as Chemistry and Physics. In the Arts Faculty itself it will be noticed that the greatest difference in examination performance is also in Mathematics 1st Course in favour of the Schoolboys.

V.—CONCLUSIONS.

The question "Do Ex-National Servicemen tend to do better at University than boys going straight there from Secondary School?" can be tentatively answered from the data obtained in the present inquiry. In Arts Ex-National Servicemen and Schoolboys with similar Entrance qualifications tend, generally speaking, to do equally well. In the Science Faculty the picture is rather different. Here both Ex-National Servicemen and Schoolboys with equivalent Entrance qualifications tend to do equally well in attaining Honours. Science Schoolboys, however, are less likely than Science Ex-National Servicemen to abandon their Degree Course completely, and are more willing to content themselves with working for an Ordinary Degree if they fail to make the necessary headway in Honours. An important factor here seems to be the comparative weakness of the Science Ex-National Serviceman's Mathematics. During National Service it would seem that he becomes rusty in this subject, and as a result is handicapped when he enters University on a Science course, since Mathematics usually figures prominently in this. A similar though less severe handicap is experienced by Arts Ex-National Service students in regard to Latin, since Latin like Mathematics tends to become rusty, and is, in effect, a compulsory subject for Arts students intending to go on to Law.

VI.—SUMMARY.

1.—Two groups of Ex-National Servicemen who entered Edinburgh University in 1949 and 1950, one group in the Faculty of Science, the other in Arts, have been made the subjects of a follow-up, and their University records compared with those of two equivalent groups of Schoolboys in the same Faculties.

2.—The results show approximately equivalent academic performance by Ex-National Servicemen and Schoolboys in the Faculty of Arts.

In the Science Faculty the results are rather different. Here both Ex-National Servicemen and Schoolboys achieve the same percentage of Honours Degrees, but the Schoolboys do better in the greater number of the remainder who achieve Ordinary Degrees, in Examination performance on their Degree Courses and in their lesser tendency than the Ex-National Servicemen to abandon Degree Courses completely.

3.—This adverse trend in the Ex-National Service students' relative performance is discussed with particular emphasis on the comparative weakness of the Ex-National Servicemen's Mathematics and the basic importance of this subject in Science courses.

BOOK REVIEWS

The Appraisal of Intelligence: A. W. HEIM. (Methuen, 1954, pp. 8+172, 12s. 6d.)

The purpose of this book is "to describe and criticize current methods of 'measuring intelligence'" and "to plead for a return to less 'objective techniques.'" It begins with a criticism of the term intelligence, as psychologists are supposed to use it, and endeavours to show how they have gone astray. Dr. Heim examines the conflicting descriptions offered by contributors to an American symposium some 33 years ago: she overlooks the fact that many of the difficulties have since been cleared up by later research. She brings the matter up to date with a further chapter on 'current definitions'; but this is confined to two—Mr. Raven's and her own. She herself maintains that psychologists should abandon any formula that 'masquerades as an exact and measurable concept,' and "give the word a meaning which is comfortably compatible with that of the layman." This, she believes, was the real intention of Binet. But, unfortunately, she only quotes him at second hand. A reference to his article will show that Binet, who of course was mainly thinking of the classification of educationally subnormal school children, makes three important distinctions. First, he distinguishes those who are lacking in intellectual capacity from those who are lacking in moral or emotional stability. Secondly, he distinguishes the lack of 'natural' capacity from the lack of acquired knowledge or 'instruction.' Thirdly, he distinguishes between what he calls 'partial aptitudes' (memory, mechanical ability, verbal ability, and the like) and the 'fundamental faculty' assumed to be common to *all* cognitive processes, from sensation up to reason—a distinction which is as old as Aristotle.

We thus reach a concept defined in three ways, namely, an innate, general, cognitive capacity. For such a concept it is convenient to have a short and simple name: Binet and his followers adopted the term 'intelligence.' The word was originally coined by Cicero to translate Aristotle's Greek label for the generic cognitive faculty; and from his day to that of Hamilton and Spencer, it was regularly used by psychological writers with this technical meaning. Like many of our older technical terms (gas, energy, soda, for example) it strayed long ago into popular parlance, and got blurred as a result. But why is the scientist now required to conform to the loose usage of the layman rather than to that of the specialist who has studied these intricate problems?

The remaining chapters are intended to present a critical review of 'objective techniques' for measuring intelligence and test-efficiency. They deal with the usual topics—I.Q., reliability and validity of tests, distribution of intelligence, and 'the approach of the factor analyst.' In the main, the author's criticisms are directed against a group of people generically termed 'psychometrists.' In spite of their claim to objective exactitude, we are told, their assessments and conclusions are constantly vitiated by 'subjective bias,' 'verbal imprecision,' and 'publicly discredited presuppositions.' These indictments are rarely supported by references¹ to specific writers or specific publications; and it is difficult to think of any competent investigators who have committed the faults that are laid at their door.

¹ One exception has caught my eye. In discussing "the alleged constancy of the I.Q.," Dr. Heim declares that "the figures given by psychometrists vary with their degree of ego-involvement . . . Rarely, if ever, are questions of differential rates and of practice openly and fully discussed. For example, Burt (*Brit. J. Educ. Psych.* XX pp. 55-61) devotes three short paragraphs to the effects of familiarity and coaching on intelligence tests." But the paper thus quoted from this *Journal* was merely a review of a book, and that is not the place for a "full discussion" of the reviewer's own researches. The interested reader would scarcely gather from Dr. Heim's remarks that I had, in fact, just published an article dealing specifically with 'The Effects of Coaching on Tests of Intelligence,' in which, besides summarizing the work of other investigators, I had given references to several detailed studies of my own, or that elsewhere I had actually printed comparative graphs giving curves for differential rates of growth.

No authoritative writer has ever regarded the mental age or the intelligence quotient as a really "scientific unit," as Dr. Heim implies. Nor do psychometrists simply "assume" that intelligence is normally distributed: by actual investigation they have proved that the distribution is *not* strictly normal, but follows a hypergeometric curve, slightly skewed for well-known reasons. They have not "ignored" the possibility that the standard deviation may vary appreciably with age: from the earliest L.C.C. surveys to those of Dr. Fraser Roberts and his colleagues, they have repeatedly shown that it does so, and devised a suitable correction. Long ago they examined the alternative procedures that Dr. Heim now puts forward—such as validating intelligence by chronological age and measuring it by percentiles—and have demonstrated the defects of each of them. They have never used 'item analysis' to measure 'reliability': Dr. Heim has probably confused 'item analysis' with 'analysing the variance of items.' And it is quite the reverse of the actual facts to suggest that 'factor analysts' in general assume "the existence in the mind of separate faculties and a God-like insight as to what these are."

No doubt a few young investigators, who might perhaps be classed as 'psychometrists' or 'factor analysts,' have been guilty of these and other careless practices; and many of the fallacies Dr. Heim has pilloried could be found in theses submitted by research-students or even in reports presented by teachers. Hence her vigorous warnings may have a welcome and beneficial effect. Certainly, in spite of—and perhaps because of—her own manifest slips and misconceptions, her book is well worth the attention of all who think of embarking on an investigation with mental or educational tests. She writes with a lively and provocative pen; and even those who disagree with her contentions will find much that is stimulating and suggestive.

CYRIL BURT.

Education and the Modern Mind: W. R. NIBLETT. (London: Faber and Faber, Ltd., 1954, pp. 155, 8s. 6d.)

This is a book of affirmations rather than of arguments. Professor Niblett finds many of the characteristics of "the modern mind" unsympathetic: it is materialistic, rationalistic, and lacking in a sense of the numinous. Professor Niblett, on the other hand, values sensibility and demands that life have meaning and purpose, even though these remain mysterious. Some will agree, others disagree with his religious belief, but all may learn from the sensitivity of feeling with which he uses it as a touchstone. This enables him to make judgments about education which may not be very novel (is it possible to say anything both new and fundamental about education?) but which are worth reflecting upon.

Readers of this journal will not expect from such a book a direct contribution to educational psychology, though it shows much intuitive insight into personality. Professor Niblett finds some scientific contributions to education stimulating, but they are those of sociology rather than psychology. Sociology emphasises adaptability of human behaviour to environmental conditions, and educationists tend to find this more encouraging than the limitations in individual capacity for learning which are reported in some branches of psychology. Yet sociology too makes the deterministic assumptions typical of a science, and Professor Niblett, valuing freedom and individuality, reminds us that scientific knowing is only one mode of experience. He reaffirms confidently the importance of value judgments, which science cannot make for us. In a wise and eloquent chapter on "Inheriting Traditions," he shows that indoctrination with values is inescapable. Perhaps he does not sufficiently recognise that it is also characteristic of the modern mind to acknowledge the importance of valuation in experience.

Professor Niblett makes thought-provoking judgments about such matters as the autonomy of schools, doctrinaire equality in education, the school curriculum and methods of training. He possibly undervalues academic and technical achievement and he has not much to say about science, but many teachers could profit by studying his sensitive approach to the teaching of literature and the arts.

R. A. C. OLIVER.

Sigmund Freud, Life and Work, Volume I—The Young Freud, 1856-1900:
 ERNEST JONES. (London: The Hogarth Press 1953, pp. 454, 27s. 6d.)

All psychologists, whatever their views as to the value of Freud's work, should surely find this an intensely interesting book. The author, the leading medical exponent of Freud's work in this country, was also himself a personal friend of Freud for forty years, and had access to a large number of private letters. Dr. Ernest Jones, as many of us know, has a notably clear style: and though he regards Freud as the outstanding genius in psychology, and dedicated the book to the daughter of "an immortal sire," he recognizes at times Freud's limitations and imperfections. In addition, he frankly states that for ten years Freud suffered from a very considerable psycho-neurosis. The conquest of this, indeed, he cites as a very great achievement.

Many readers will learn for the first time of the great difficulties which Freud had to face at the earliest stage of his career—financial problems, and social relations within his own family, and that of his future wife. With these and with more professional difficulties he struggled manfully; his courage and persistence were, indeed, astonishing; but in his relations with professional friends we see already signs of a strong resentment of criticism of his own views.

Dr. Jones constantly tries to explain the attitudes of Freud, and events in his relations with others, by Freud's own mature views as to unconscious influences. Here he is much less successful, and sometimes an event which seems to have a perfectly simple explanation is interpreted in a quite unnecessary way; thus Freud was dissatisfied with one of his chiefs—Claus, and greatly admired one Brücke. Now, writes Dr. Jones: "Claus was twenty years older than Freud, while Brücke was forty years his senior. These differences in age correspond exactly with on the one hand, that between Freud and his half-brother, the imagined rival with his mother in early childhood, and on the other hand that between Freud and his omniscient and beloved father."

The account of Freud's self-analysis leading to the discovery that, when a little boy he had unconsciously hated his father, the present reviewer at least, also finds quite unconvincing. Nevertheless, the book makes excellent reading. It will no doubt be illuminating to different readers in very different ways. C.W.V.

Marriage Failures and the Children: CLAUDE MULLINS. (The Epworth Press, London, 5s., pp. 60.)

The author of this book is well-known as a magistrate of long experience. In fifteen years on the bench he has dealt with about 5,000 matrimonial cases. His sense of responsibility and his deep concern for the welfare of the children of parents who are divorced, are evident throughout the book. He makes a strong case for increasing state expenditure on advice and guidance for married couples when divorce is contemplated, especially in view of the heavy expenditure now involved in giving legal aid for divorce. Lawyers, he argues, are not necessarily the best counsellors on the special problems of married life. "In the three years during which £12,000 was paid by the State for Marriage Guidance, no less than £1,200,000 was paid for free and assisted legal aid, mostly for divorce cases."

Mr. Mullins contests the view that it is usually better for a child to live permanently with one parent rather than with two who are always quarrelling. Here, however, he offers less convincing evidence: some recent investigations suggest that in itself the broken home is not necessarily in itself as bad an influence on the child as some have thought.

Still, there are no doubt many couples seeking divorce, who selfishly give too little thought to the possible effects on the children if young, and Mr. Mullins' plea is they should at least seek to continue their life together until the youngest child is seventeen. The age of seventeen may be unnecessarily high in view of the author's explicit acceptance of the idea that "The fundamental traits of character are settled when a child reaches the age of seven." It seems, however, that he thinks all the children should be settled in an occupation before the parents separate.

The Abilities of Babies: RUTH GRIFFITHS. (Univ. of London Press, 1954, pp. 229, 20s.)

On the basis of her own preliminary observations made on babies under the age of two and the records of earlier observers and testers, especially Gesell, Dr. Griffiths has drawn up scales of tests or, as most might better be called, items of behaviour, assigned to each of the months from 1 to 24.

Gesell classified his items and tests into four groups—Motor, Language, Adaptive and Personal—Social. Dr. Griffiths, dissatisfied especially with Gesell's placing of some items under 'Adaptive behaviour,' has grouped her items under five categories, viz.: Locomotion, Personal—Social, Hearing and Speed, Hand and Eye, Performance. She lists fifty-two items in each scale for the twenty-four months period. These items have been standardised on 552 infants, the occupations of whose fathers show a fair scatter from professional to semi-skilled, but with very few labourers and no rural workers.

A great amount of work has gone to the preparation of these scales and Dr. Griffiths must be commended for her patient industry. She rightly stresses the advantage of having the same tester for all ages and all lists.

I welcome warmly also Dr. Griffiths' recognition that at these early ages no single testing can be relied on: indeed, in view of the great variability of mood and momentary interest of a child of under two or three, three or four sessions would be the minimum even after the child has become quite accustomed to the tester.

Dr. Griffiths's assignment of a given item to a specific month is hazardous as the average number of babies of a given month was only twenty-five; and indeed for one month there were only sixteen, and Dr. Griffiths states that some of the children included in the results for the second year were already included in the first year, though there was no retesting in less than six months. The various socio-economic levels also might not have been fairly represented for each month.

The chapters on standardisation and frequency curves show similar results to averages and scatters gained with recognised Intelligence Tests at later ages; but as to 'validity' Dr. Griffiths only offers some results gained by retesting a group of infants with her own series. Her scales, however, do provide a valuable basis for follow-up studies; and it is to be hoped that she will be able to retest a substantial number of her infants with recognised standard tests at say 5 and 10 years. Such a research might lead to the selection of certain of her present many items in each scale, as being of much greater prognostic value than others, and the elimination of many as of little use.

Special mention should be made of the value of the arrangement of the scales for children with various forms of handicaps, e.g., for deaf, blind or spastic babies.
C.W.V.

Geometry of Mental Measurement: SIR GODFREY H. THOMSON. (Univ. of London Press, 1954.)

There is now nearly a complete generation of research students in psychology and education who must have received much of their instruction on factorial techniques from Godfrey Thomson's *Factorial Analysis of Human Ability*.

The book under review also deals with the same problem, but is restricted to mathematical discussions and suggestions about factorial analysis. The material is clearly and precisely presented, which by its nature could give difficulty to students having "only a very meagre mathematical equipment." Godfrey Thomson, however, has demonstrated again how lucidly he can expound this material. He uses the geometrical model of factor analysis and this is the one which occurs most readily to the beginner. For such a person, the present volume could act as an intermediary between the accounts of the arithmetical processes of factor analysis set out in the body of the earlier *Factorial Analysis*, and the concise matrix treatment provided in the arithmetical appendix of the same volume. In fact, a good exercise for the earnest research student would be to translate the contents of *The Geometry of Mental Measurement* into the terms of the mathematical appendix of the earlier volume.

Thomson deals with all the important problems of the geometrical interpretation of factors and introduces the reader to the field of prediction and estimation. He would do a further service to this branch of psychometrics if he were to do the same for problems of maximum prediction of multiple criteria. In all, this clear little book should be read by every student of psychometrics and studied with great care by those writing theses on factorial research.

E. A. PEEL.

The Doctrines of the Great Educators: R. R. RUSK. (Macmillan and Co., 10s. 6d. pp. 311.)

This is a revised edition of a book first published as long ago as 1918. It is a testimony to its usefulness that reprints have been required frequently. In this second edition some chapters have been re-written and a useful chapter on Dewey added.

OTHER PUBLICATIONS RECEIVED

The mention of a book in this list does not preclude a later review.

- Psychology*: W. J. PITT and J. A. GOLDBERG. (McGraw-Hill, pp. 413, 38s.)
- The Origins and History of Consciousness*: E. NEUMANN. (Routledge and Kegan Paul, pp. 493, 30s.)
- From Loneliness to Fellowship*: W. AAREK. (Allen and Unwin, pp. 79, 3s. 6d.)
- La Cybernétique*: G. T. GUILBAUD. (Presses Univ. de France, pp. 135, no price given.)
- Developments in the Rorschach Technique*: B. KLOPPER, M. D. AINSWORTH, W. G. KLOPPER, and R. R. HOLT. (Harrap, pp. 726, 42s.)
- The Teaching of Arabic Numerals*: G. G. NEILL WRIGHT. (Univ. of London Press, pp. 45, 2s. 6d.)
- The Durham Research Review* (The Research Publication of the Institute of Education, Univ. of Durham): Edited by F. V. SMITH. (Univ. of London Press, pp. 62, 5s.)
- Speech in the Elementary School*: M. OGILVIE. (McGraw-Hill, pp. 318, 38s.)
- Personnel Problems of School Administrators*: Clarence A. Weber. (McGraw-Hill pp. 378, 40s.)
- Research Methods in the Behavioural Sciences*: Edit. by L. FESTINGER and D. KATZ. (Staples Press, pp. 660, 50s.)
- The Prediction of Student-Teaching Success from Personality Inventories*: F. T. TYLER. (Univ. of California Press, pp. 313, \$1.25.)
- Report—First International Congress for Teaching Educational Sciences in Universities* (Ghent, 1953), *Proceedings and Reports on the Teaching of Educational Sciences in Western Europe*: R. L. PLANCKE and R. VERBIST. (Ghent, pp. 248, no price given.)
- Crime and the Services*: JOHN SPENCER. (Routledge and Kegan Paul, pp. 306, 28s.)
- Isaac and Oedipus*: ERICH WELLISCH. (Routledge and Kegan Paul, pp. 131, 15s.)
- Techniques of Counselling*: JANE WATERS. (McGraw-Hill Book Co., pp. 384, 38s.)
- The Development of Personality—Vol. 17*: Collected Works of C. G. JUNG; Edited by SIR HERBERT REED, M. FORDHAM and G. ADLER. (Routledge and Kegan Paul, pp. 235, 25s.)
- Pain and Other Problems*: J. C. WORDSWORTH. (Allen and Unwin, pp. 159, 12s. 6d.)
- The Collected Papers of Otto Fenichel, First Series*: Edited by H. FENICHEL and D. RAPAPORT. (Routledge and Kegan Paul, pp. 408, 35s.)



SIR GODFREY THOMSON

SIR GODFREY THOMSON

1881-1955

SIR GODFREY THOMSON was born on 27th March, 1881, and died on 9th February, 1955. He was educated at Rutherford College, Newcastle, and at the University of Durham, where he graduated B.Sc. with distinction in Physics and Mathematics. He was later appointed Pemberton Research Fellow of the University of Durham and went to Strasbourg to do experimental research, under Ferdinand Braun. At about this time his interest turned from Mathematics to Education, and he became lecturer and later Professor of Education at Armstrong College, Newcastle. In 1925 he was appointed to the joint posts of Professor of Education in the University of Edinburgh and Director of Studies at Moray House—an appointment which he held until his retirement in 1951.

He was President of the British Psychological Society in 1945-46, and President of the Psychological Section of the British Association in 1949. He was a member of the editorial board of the *British Journal of Educational Psychology* for many years and joint editor (with Sir Cyril Burt) of the *British Journal of Psychology, Statistical Section*, from 1947 until 1952. He was honoured by his own University of Durham and by learned societies in America and in Sweden. The Polish Government in exile admitted him to the order of Polonia Restituta and his own country, recognising his contribution to colonial education as well as to education at home, honoured him with the dignity of knighthood in 1949.

In a sense, the whole of his work in the field of educational and psychological research has been an attempt to bring mathematical exactitude into psychological experiment and theorising. It can be divided into three parts, not of course, quite separate, and over-lapping in time. These parts are concerned with (a) fitting psychophysical curves; (b) the social and geographical distribution of intelligence and the influence of the differential birthrate; and (c) the factorial analysis of ability.

His interest in psychophysics began in 1911 and he examined the mathematical bases of the various methods of estimating absolute and difference thresholds of sensation. Most of the mathematics concerned with the significance or accuracy of these estimates he found in a very shaky condition and he endeavoured to make it more exact.

Most of his work in large-scale intelligence surveys developed out of a desire to give an equal educational chance to children in different classes of society and in different districts, and throughout his life he laboured unceasingly to hold open the gates of higher education to all able children no matter how humble their birth. To this end he developed the tests of intelligence and attainment which, when he came to Edinburgh, became known as the Moray House Tests and which in his later years sold in millions. With characteristic integrity he devoted the whole profits from these tests to the promotion of educational research and to the provision of bursaries for able students.

The most extensive intelligence surveys to which he contributed all his knowledge and experience were the two Scottish Mental Surveys in each of which every 11-year-old child in Scotland was tested. At the time of his death, he was actively concerned in the follow up of a sample of the children tested in the 1947 survey.

The third branch of his work, however, on the factorial analysis of human ability was the one which interested him most strongly and was also the one

in which he carried out the most original work. His controversies with Spearman and Burt in this country, and Thurstone in America, are well-known to those who have worked in the field. In his book *The Factorial Analysis of Human Ability*, he has described at considerable length (and in some cases more clearly than they have expressed it themselves) the work of Spearman and Thurstone, as well as his own contribution. The main theorem for which he claims authorship is the discovery that a large number of small independent causal influences of which samples come into play in different activities of the mind act exactly as if there was one general cause common to all of these activities, and thereafter specific causes each confined to one activity only. If the number of such causal influences is large enough, if there are no linkages between them and if each causal influence is "all or none" in action (i.e., comes into action with its full force or does not come in at all), then the matrix of correlations between the activities or 'tests,' taken in pairs, can be reduced as closely as we wish to rank one (apart from diagonal entries).

Articles on all of these topics have appeared at intervals in the British and American psychological journals and in other scientific journals.

It is too early yet to assess Sir Godfrey's true contribution to education, but many of his students will feel that his main contribution was made through his teaching. He was a most lucid lecturer possessing a remarkable command of language and the great gift of selecting precisely the best way of expressing an idea in order that it would carry most force. His assessment of the views of some of the great educators of our times will be remembered by many of his students long after other memories of their course have faded. It was his constant endeavour to know each one of his students personally and he took a very lively interest in their welfare, even long after they had left Moray House.

The students who were closest to him, however, were those who were reading for the degree of Bachelor of Education, particularly in their final year. There his skill as a lecturer was shown at its best; there his extraordinary capacity for interesting an intelligent but mathematically unqualified audience in the intricacies of educational statistics was used to its fullest extent.

It was those students who benefited most from his own very clear insight into the part that educational research can play and from his very careful assessment of its powers and limitations. These students he not only equipped with the techniques for carrying out research but to each he also conveyed a little of the inspiration which he himself possessed.

In a period when the demand for well qualified men and women was great in the field of educational administration, in educational clinics and in the universities, he (with the late James Drever) gave to these fields a large number of young people possessing not only the necessary academic qualifications, but also the sense of responsibility which is necessary for all whose work deals with human beings and with a balanced outlook on the problems of education.

In the final reckoning it may well be that his students will be his finest memorial.

JOHN SUTHERLAND.

ENVIRONMENTAL AND OTHER CHARACTERISTICS DETERMINING ATTAINMENT IN PRIMARY SCHOOLS*

By LESLIE C. D. KEMP

(*Armidale Teachers' College, and University of New England, Armidale, N.S.W., Australia.*)

I.—Introduction. II.—Scope and Methods. III.—Results. IV.—Summary.
V.—References.

I.—INTRODUCTION.

THERE have been many investigations into factors associated with the educational progress of school children. Burt (1917) showed very clearly that intelligence was the predominant constituent of general educational ability, as many schoolmasters had suspected long before. His conclusion was re-inforced in a more recent study. (Burt, 1939.) He also suggested that memory, interest and industry were other important factors in educational progress. This again, of course, was a confirmation of views held much earlier by schoolmasters. Some similar findings were later revealed in Alexander's research (1935). Fleming's review of researches on the relations between socio-economic status and test performance (1943) confirmed the view that the social and economic level of a child's environment was associated with the level achieved in tests of intelligence and school attainments. In a more recent investigation in London, Campbell (1952) claimed that cultural level was apparently an important factor in determining success in secondary schools. In all these studies the factors may be regarded as individual ones since they relate to the attainment of individual children. It is likely that characteristics of a school are also important in determining attainment. Teachers, as well as educational theorists and administrators, are naturally concerned with such things as the school buildings and amenities, the size of classes, and methods of school organization and of teaching. But on all these points far less attention has been focussed than on the individual psychological and physical aspects mentioned above.

During the past fifty years a number of rough investigations have been conducted into the relationship between achievement in schools (e.g. as assessed by scholarships won) and the type of school or the area where it was situated. But the chief "school" (as distinct from "individual") factor which seems so far to have been investigated is *progressiveness-orthodoxy*. Thorndike (1941) compared various aspects of behaviour in children who had attended both progressive and orthodox elementary schools, and found slight differences. In this country, Gardner has investigated the attainment of children in orthodox and progressive infant schools (1942) and has also compared children in the junior school who had previously attended the two contrasted types of infant school (1950).

Ever since the Board of Education urged the introduction of activity

* This article is based upon an investigation carried out at the University of London Institute of Education under the guidance of Professor P. E. Vernon, for whose generous advice and helpful criticism I am very thankful. I acknowledge with deep gratitude the financial assistance provided, in connection with the research, by the Imperial Relations Trust, the Leon Bequest Committee, and the Central Research Funds Committee of the University of London. I should also like to thank the London County Council for permission to carry out the work in its schools, and the various Divisional Officers, Inspectors, Head Masters and Head Mistresses for their co-operation. My sincere thanks are due to Sir Cyril Burt for his comments and advice in connection with this article.

work into the primary school in 1931* there has been a noticeable trend towards a more progressive school organization. At the same time, many teachers have remained almost belligerently, or at least defiantly, orthodox. Activity methods have been much more widely employed in infant schools, although many junior schools are now organized along these lines. Miss Gardner's first investigation (Gardner, 1942) was concerned with the relative effectiveness of the progressive and orthodox organizations at the infant level. In a few specially selected schools she found that, far from being inferior to the children in the orthodox schools in some of the things which are regarded as important outcomes of infant school education, the children in the progressive schools tended to be at least equal, and in some respects superior, to those in the orthodox ones. In her second study (1950) she made comparisons, again in just a few schools, between junior school children aged 9 and 10 who had attended a progressive infant school, and those of the same age who had been to an orthodox one. The junior schools used were "good junior schools of a traditional type." Again the evidence seemed to favour the children who had attended the progressive infant schools. It must be stressed that not only was the number of schools very small, but they were specially selected; and the classical experimental design of matched groups was used with an examination of the significance of differences between mean scores by a method which is open to criticism. Moreover, in so far as the schools were specially chosen as "special cases," the results are not necessarily the same as might be found in a larger group of schools of varying degrees of progressiveness.

It is well known to those who work in them that each school possesses a characteristic tone or morale, which is no doubt related to the qualities of the Head and Staff as well as to other factors, such as the type of community which the school serves. Cattell has recently written about the characteristics of groups (Cattell, 1948) and has suggested that it might be possible to study the personality (or *syntality*) of groups much as is done for individual personalities. It was thought that both what Cattell calls *syntality traits* and *population traits* might be investigated, among primary schools, for their connection with attainment. It was decided to secure measures of a large number of such traits for fifty junior mixed schools, and to investigate their interrelations by means of modern statistical techniques. For a number of reasons it was thought desirable to confine the investigation to schools in two educational divisions of the London County Council. While the schools used could not be regarded as a random sample of London's primary schools they were to a considerable extent representative of schools under that authority.

II.—SCORE AND METHODS.

An attempt was made to assess what may loosely be called the Head's educational philosophy through its manifestations in the school *atmosphere* and *type of organization*. These assessments were based upon observations made by the author while in each school. The period spent in each school was devoted to testing pupils, talking with the Head and Staff, and observing the conduct of the school generally. This period was one full day, which, from some points of view, was perhaps inadequate, and a longer period would no doubt have permitted slightly more valid and reliable judgments. But time was limited to this period by various factors beyond the author's control.

Atmosphere varied from that found in the cold, silent, machine-like

* Report of the Consultative Committee of the Board of Education of Great Britain and Northern Ireland on "The Primary School," where it was urged that "the curriculum of the primary school must be thought of in terms of activity and experience, rather than of knowledge to be acquired and facts to be stored."

school with unsmiling children and stern teachers, on the one hand, to that observed in the warm, happy and vital school, on the other. Ratings were made on a five-point scale.

Type of organization varies from the very formal, rigid and orthodox to the most informal, free and progressive, with a curriculum organized through activities and experiences related to the needs and interests of the children. To assess type of organization, a descriptive rating scale was constructed with five typical descriptions varying from the extremely progressive to the severely orthodox. This scale was used independently by both the author and the Inspector of schools. The two sets of judgments of fifty schools, made quite independently, were found to be highly correlated: the coefficient was .78.

While atmosphere and type of organization may be thought to be predominantly due to the Head's philosophy and methods, the attitudes of the staff also exert a strong influence on the behaviour of the pupils. Contact with the staff was made in both individual conversations during visits to schools and in the group contexts in the staff room. An assessment was made, in the form of a rating, of what I have called *staff spirit*. This is the degree of happiness and contentedness and interest shown by the staff as a group.

A second group of characteristics is related to the physical features of the school itself, and of these the actual size or *enrolment* of the school and the mean size of classes were thought to be worth investigating.

A third group includes features of the site and buildings where the school is housed. Here were included such things as *age of the building*, its *interior state of repair and attractiveness*, the *size of the playground*, and the adequacy and cleanliness of the *sanitary facilities*.

Finally, there is the type of neighbourhood where the school is situated and which it serves. I have called this the catchment area, and it was rated for *cleanliness and respectability*. This was based upon observations made in visiting the school for testing and again during a further visit to the neighbourhood at a later date. The things taken into account here were: type and age of houses, their state of repair and attractiveness, presence of front garden, general appearance and tidiness of streets, other kinds of buildings, and general appearance of the people in the street.

The *economic status* of the catchment areas was assessed by the use of the recently developed J-Index.* For each school, the catchment area embraced the streets within a circle of approximately one-third of a mile radius, with the school as centre. From the Registers of Electors in the local libraries the figures were ascertained for each of these streets, and from these the J-Index was calculated.

Inquiries about the occupations of fathers were not permitted, so a rough estimate of the *paternal occupational level* for each school was made from the Headmaster's knowledge of the children's home backgrounds. *Size of family*, percentage of *mothers working*, percentage of children *learning a musical instrument*, and percentage of *homes possessing a car, radio, television set, or piano* were determined from a questionnaire submitted to the children.

* Because the more traditional methods of estimating the economic status of a district (based upon the rateable value of property) suffer from the disadvantage that they can only be used for the administrative districts for which such figures are available, the J-Index was developed as an alternative means of assessment. This index is based upon the percentage of electors who are liable for jury service, and may be used for any defined area at all. People so liable are indicated in the Register of Electors by the suffix "J". Further information about the J-Index may be found in Gray, *et al.* (1951). Its validity is quite high, when judged against the other accepted indices of economic status, and its reliability in this investigation was estimated—*via* communality—as at least .70.

In addition to the foregoing characteristics, the *abilities* and *attainments* of the children had to be assessed. Other data included *regularity of attendance* (as indicated by mean percentage of children attending during three successive months) and *height* (mean height of children tested). About one-third of the schools used were voluntary schools, and exactly one half of them had male Heads. Thus *type of school control* (county and voluntary) and *sex of Head* were included as two other variables which might possibly be related to level of attainment. An assessment was also attempted for children's *interest in school*. They were presented with a list of ten activities of varying degrees of attractiveness and interest for children and asked to indicate their order of preference for doing these things if they had a spare day on which they had to select one of these activities. One of the ten was simply "At school." The mean rating given to this item at a school was used as a measure for interest in school.

Because the time available for visiting schools to collect the data was limited to 15 weeks and because the attention and interest of eleven-year-old children (particularly the duller ones) is not sustained, a programme of short and varied tests was devised. It was realised that on the whole the reliability of shorter tests tends to be lower than that of longer ones, but it was thought possible to secure a compromise between the demands of time and of pupil interest on the one hand and of reliability on the other.

Two measures of intelligence were used. Vernon's Abstraction Test has been widely used in Britain, and its validity and reliability are very satisfactory. It is an excellent illustration of Spearman's education of relations and correlates; its correlation with intelligence is known to be quite high: with verbal ability and education, low. The adaptation used in this investigation contained twenty items, with a ten minute time limit.

As a Non-Verbal test of intelligence, a five-minute test was devised, with twenty items, based upon one used some years ago by an L.E.A. in connection with selection for secondary schools.

A test similar to, though shorter than, the Watts-Vernon Silent Reading Test was used to assess *reading comprehension ability*. The items were straightforward and based upon every-day situations.

Tests of *Mechanical Arithmetic* and *Problem Arithmetic* were constructed; the former was based upon the four operations with number and money, including addition and subtraction of fractions. The Problems Test contained twenty-two graded problems involving various operations, and was based upon tests originally devised by Burt and Thomson.

A modified form of Burt's Graded Word Spelling Test, containing twenty-one words ranging from the seven-year to the thirteen-year level, was used to test *spelling*.

Writing was judged for both *speed* and *quality*. The former was gauged by the number of words the children could write in two minutes; the latter was judged from the spelling scripts using a quality scale constructed therefrom.

The assessment of skill in English *composition* is not easy to accomplish quickly. An idea of Watts (1944) was used, and the children were required to write two three-sentence compositions. Their quality was rated on a five-point scale, and these were combined to provide the final assessment.

Finally, a *General Information Test* was compiled based upon the research done by Manley (1951). The final form of this test contained forty varied questions with an eight-minute time limit.

A number of other assessments of the children were made. These included ratings for *sociability*, *co-operativeness* and *resourcefulness*, as well as for their general *manageability* in the classroom, and their playground behaviour. The

first three were rated upon the observed behaviour during the time spent in each school, and particularly during a piece of impromptu dramatisation which was done at the close of the testing programme. Manageability ratings were based upon the entire day's experience with the children, and *general pressure of playground activity* was based upon observed playground behaviour at one of the recesses.

III.—RESULTS.

The school scores were based upon the scores obtained by sample groups of children in each school. In all schools the sample came from the children in the final junior year. Where the total enrolment in this final year was not more than about thirty, all were tested. Where it was larger, the youngest and oldest were eliminated, month by month, from the total group till a group of about thirty remained. The schools were all mixed, and the sample groups contained roughly the same number of each sex. Mean scores were calculated for boys and girls separately, and these were then averaged to determine the school score.)

The testing programme was spread, of necessity, over a period of five months. Since groups with approximately the same mean date of birth were used, it was clear that there was a slight increase in mean age over the period of testing. An investigation into the effect of this variation in age from group to group on the scores obtained showed that there was no justification for rejecting the null hypothesis that the groups did not differ in test scores merely because of differing age.

Scores were thus determined for the following forty-two variables.

- | | |
|---|--|
| 1.—Nature of district as indicated by the J-Index of Economic status. | 19.—Handwriting quality. |
| 2.—Paternal occupational level. | 20.—Composition. |
| 3.—Percentage of children learning a musical instrument. | 21.—Information Test. |
| 4.—Percentage of children whose mothers do not go to work. | 22.—Abstraction Test. |
| 5.—Percentage of homes with a piano. | 23.—Non-Verbal Test. |
| 6.—Percentage of homes with radio. | 24.—Resourcefulness. |
| 7.—Percentage of homes with television. | 25.—Co-operativeness. |
| 8.—Percentage of homes with a car. | 26.—Sociability. |
| 9.—Percentage of homes with telephone. | 27.—Playground behaviour. |
| 10.—Size of family (Positive score for "smallness"). | 28.—Manageability of children. |
| 11.—Cleanliness and respectability of neighbourhood. | 29.—School attendance. |
| 12.—Area of open spaces in catchment area. | 30.—Mean height. |
| 13.—Interest in school. | 31.—Sex of Head. |
| 14.—Mechanical arithmetic. | 32.—Staff spirit. |
| 15.—Arithmetical problems. | 33.—School atmosphere. |
| 16.—Silent Reading. | 34.—Enrolment. |
| 17.—Spelling. | 35.—Mean class size. |
| 18.—Handwriting speed. | 36.—Progressiveness, as judged by the author. |
| | 37.—Progressiveness, as judged by Inspector. |
| | 38.—Age of building. |
| | 39.—Interior state of repair and attractiveness. |
| | 40.—Sanitary facilities. |
| | 41.—Playground space per pupil. |
| | 42.—County or voluntary school. |

This multitude of assessments was next reduced to a smaller number of descriptive variables partly on the basis of common sense and partly by the use of factorial and other analyses. It was found that variables 1, 2, 9, 10, and 11 were closely related to one another, and that collectively they would provide a score for what may be called *socio-economic status*. Among variables 14 to 21 a considerable amount of agreement was found in school scores. Schools whose pupils were good at any one subject were found to be good at the others. It was thought justifiable, however, on the basis of the statistical evidence, to use two criteria of attainment. The former included Arithmetic Problems, Reading and General Information; the latter, the remaining subjects. Among the scores for various aspects of the children's behaviour two useful groupings were discovered. The first contained resourcefulness, co-operation and sociability, and the second, school atmosphere, manageability, playground behaviour and regularity of attendance. The former cluster of qualities was called *adjustment*, and the latter school *morale*. A criterion of *intelligence* level for each school was set up by combining the scores for variables 22 and 23. The Inspectors' and the author's ratings were combined together to provide a criterion of *progressiveness*. Variables 38, 39 and 40 were amalgamated into a single score for *school building*. To these eight criteria were added the remaining eight from the above list to form a list of sixteen variables composed of various criteria of "environmental and other characteristics" and "attainment." The intercorrelations among these sixteen variables are given in Table 1.

The highest coefficient in Table 1 is .73 between attainment (Comprehension) and intelligence. Thus, when schools are considered as units, rather than individual children, it is still found that the greatest single factor determining level of attainment is intelligence. More than half of the variance in attainment in reading, problem arithmetic and general information is accounted for by this factor. However, it must be noted that socio-economic status is correlated very significantly with both intelligence (.52) and this kind of attainment (.56). When the former is partialled out the correlation drops to .62. If intelligence is held constant, the correlation between socio-economic status and attainment drops to .30. The influence of socio-economic status on attainment appears thus to be much less powerful than that of intelligence.

In the rote subjects, the same picture emerges. With intelligence the correlation is .61; with socio-economic status, .47. The former drops to .50 when status is held constant; and the latter to .23 when intelligence is held constant. Again it is clear that intelligence is the dominant factor determining attainment in mechanical arithmetic, spelling, writing quality and simple composition.

The multiple correlations between a combination of intelligence and socio-economic status and each kind of attainment are found to be .76 and .65. Thus 57 per cent. of the variance in comprehension subjects and 42 per cent. in rote subjects is accounted for by these two factors.

Both kinds of attainment in school subjects are related quite significantly to school morale, ($r = .51$ and $.48$). It may be that schools with better morale reach a higher standard of attainment partly for this reason, or that factors which lead to better achievement also lead to better morale. It is noticeable that both intelligence and socio-economic status are also significantly correlated with morale ($r = .45$ and $.34$), so that it is possible that these two factors are important predisposing causes in the formation of school morale just as they are in the production of better achievement; both morale and achievement requiring, of course, the appropriate kind of teaching for their realisation.

TABLE I

INTERCORRELATIONS AMONG THE SIXTEEN VARIABLES. POSITIVE SCORES HAVE BEEN DESCRIBED IN THE FIRST COLUMN.

	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)	(x)	(xi)	(xii)	(xiii)	(xiv)	(xv)	(xvi)
(i) (Upper Class)	—	52	56	47	32	34	18	11	15	-10	-05	-05	-27	16	33	29
(ii) (Intelligent)	52	—	73	61	43	45	18	24	10	-12	-07	16	-17	02	13	25
(iii) (High comp. attain.)	56	73	—	71	20	51	30	17	16	-21	09	14	-06	12	03	11
(iv) (High rote attain.)	47	61	71	—	25	48	39	11	13	-14	01	10	-08	00	08	05
(v) (Keen Interest)	32	43	20	25	—	17	-14	-03	-07	-04	-22	-11	-39	17	08	17
(vi) (Good morale)	34	45	51	48	17	—	00	-09	06	23	-09	06	-20	14	39	07
(vii) (Large enrolment)	18	18	30	39	-14	00	—	36	28	07	48	24	47	-32	00	13
(viii) (Male Head)	11	24	17	11	-03	-09	36	—	31	00	00	00	08	00	-07	14
(ix) (Progressive school)	15	10	16	13	-07	06	28	31	—	14	16	32	18	-03	16	18
(x) (Large playground)	10	-12	-21	14	04	23	07	00	14	—	24	11	44	00	-03	23
(xi) (New building)	-05	-07	09	01	-22	-09	48	00	16	24	—	23	58	14	-04	26
(xii) (Happy staff)	05	16	14	10	11	06	24	00	32	11	23	—	25	05	05	10
(xiii) (County school)	27	-17	06	08	39	20	47	08	18	44	58	25	—	21	-27	-04
(xiv) (Small classes)	16	02	12	00	17	14	-32	00	03	00	-14	05	-21	—	11	11
(xv) (Good adjustment)	33	13	03	08	08	39	00	-07	16	-03	-04	05	-27	11	—	35
(xvi) (Tall children)	29	25	11	05	17	07	13	14	18	23	26	10	-04	11	35	—

NOTE.—All decimal points and plus signs have been omitted.

The correlation between morale and (comprehension) attainment drops from $\cdot51$ to $\cdot43$ when socio-economic status is held constant, and to $\cdot30$ when intelligence is held constant. If both are partialled out it drops to $\cdot23$, which is below the 5 per cent. level of significance. The same kind of thing happens with rote attainment. It drops from $\cdot48$ to $\cdot40$ or to $\cdot30$ according as socio-economic status or intelligence is partialled out; and to $\cdot25$ when both are held constant.

The only other characteristic which is significantly related to attainment in the formal subjects is enrolment. ($r = \cdot30$ and $\cdot39$.) It is plausible that Heads seeking promotion from smaller to larger schools would tend to be the more enthusiastic and ambitious ones, and that those with reputations for efficiency would be more likely to be selected. If this sort of reasoning is correct, it seems equally likely that these Heads would gather around them staffs who would be keen, enthusiastic, and efficient. Large schools tend to be housed in newer and more attractive buildings, to be "county" schools, to have a man in charge, to be more progressive and possibly to have a happier staff. The last three of these characteristics have positive, though slight, correlations with attainment. The multiple correlation between these three and attainment is $\cdot24$.

Another variable which is almost significantly related to both kinds of attainment is *interest in school* ($r = \cdot20$ and $\cdot25$). Children's interest is no doubt fostered as much by their success as is success by interest in school. Children in schools where the mean rating for interest in school is higher tend to be more intelligent ($r = \cdot43$), and to come from homes of higher socio-economic status ($r = \cdot32$); and the schools tend to be "voluntary" schools ($r = \cdot39$), and to be housed in older and less attractive buildings ($r = -\cdot22$ with new buildings). This is an interesting finding. Although the actual correlation between newness of building and interest in school is below the 5 per cent. level of significance, the clustering of interest with small orthodox schools in old buildings under denominational control is quite clear. This is relevant to the frequently heard suggestion that older buildings must be demolished and that new and attractive buildings of modern design are essential. So far as the children's interest and happiness at school is concerned, the effect of newer buildings appears to be doubtful. It is also worth noting that interest in school appears to be independent of whether the school is progressive or orthodox in organization.

Just as it has been suggested that morale may be thought of as a kind of attainment, so with *interest in school*. The correlations between interest and attainment are reduced to zero when the effects of intelligence and socio-economic status are partialled out. Interest is apparently determined in some measure by intelligence and socio-economic status, but it is also significantly related to "voluntary" control ($r = \cdot39$).

Adjustment is not significantly related to attainment ($r = \cdot03$ and $\cdot08$). On the other hand it is correlated significantly with morale ($\cdot39$), tallness ($\cdot35$), socio-economic status ($\cdot33$) and "voluntary" schools ($\cdot27$). Each of these four variables is correlated positively with attainment in the formal school subjects, and with each other.

The high correlation between morale and adjustment may arise partly because they are both based upon the author's own ratings. Both morale and adjustment are significantly related to socio-economic status. The multiple correlation between adjustment and a combination of morale and socio-economic status is $\cdot42$. This almost 20 per cent. of the variance in adjustment is accounted for by these two factors.

Two variables which have not emerged as important determinants of any kind of attainment are *progressiveness* and *class size*. Table 1 shows that the correlations between progressiveness and the two types of attainment are both below the 5 per cent. level of significance. The only significant correlations with progressiveness are those with staff spirit ($\cdot 32$) and enrolment ($\cdot 28$). Other small (though statistically insignificant) positive correlations with progressiveness are those with intelligence ($\cdot 10$), socio-economic status ($\cdot 15$), school building ($\cdot 16$), large playground ($\cdot 14$), "county" control ($\cdot 18$), male Head ($\cdot 16$) and tallness ($\cdot 18$). Thus with progressiveness there is a loose clustering of happy staff spirit, large school, "county" control and male Head. All of them are negatively, though insignificantly (except for "county" control) correlated with interest in school, which also shows a slight negative correlation with progressiveness itself. Morale seems to be quite unrelated to all these factors and to be independent of progressiveness or orthodoxy. Adjustment is even less closely related to progressiveness than is attainment in the formal school subjects.

Nevertheless there is a slight tendency for progressiveness and formal school achievement to be positively correlated. The correlations are small ($\cdot 16$ and $\cdot 13$) and are reduced to an even smaller magnitude if enrolment is held constant. Thus the effect of progressiveness on attainment is good rather than bad; but by no means marked. One can say that progressiveness is much less potent in its influence on junior school attainment than either enrolment or morale. There is no evidence in this investigation that progressiveness is harmful in its effect on attainment, nor that it is particularly helpful. This is much like the conclusion reached by Miss Gardner about the long term effects in the early junior years of the progressive infant school training.

The correlations between small class size and attainment are also very low. Within the range of mean class size in the schools used (25 to 46) the size of class appears to be unrelated in any statistically significant way to attainment. However, it was noted that smaller classes tended to be found in smaller schools ($r = \cdot 32$); and it has been found that enrolment is related to attainment. When enrolment is held constant, the partial correlation between class size and attainment rises to $\cdot 24$ (for comprehension) and $\cdot 14$ (for rote). Only the former coefficient approaches the 5 per cent. level of significance. The mean size of classes in the five schools with the highest attainment in the comprehension subjects were: 29, 37, 39, 41, and 44. In the rote subjects they were: 37, 38, 39, 40, and 44. If anything, there is less relationship between class size and rote attainment than between class size and comprehension attainment; and rote attainment than between class size and comprehension attainment; which is contrary to what one might expect on *a priori* grounds, since it is thought that smaller classes are needed for the "drill" subjects so that more individual and small-group tuition may be given. The present investigation provides no evidence to show that large classes necessarily have a deleterious effect on school attainment. It is, of course, possible that this tendency might actually exist, but has been here obscured by the fact that the large classes, occurring in the larger schools, are likely to be more homogeneous. But these are not points with which this investigation was directly concerned.

The effects of the various factors on attainment were also expressed in terms of E.Qs. Analysis of variance was used to test the significance of the difference in E.Q. between the upper and lower quarters of the distribution of scores for each of the conditions assessed. Table 2 shows a summary of this step. The scores for comprehension and rote attainment were combined to provide one single measure for attainment, and of the fourteen characteristics used, the six with the greatest differences between upper and lower groups are

shown. For intelligence, socio-economic status and morale the differences are significant at the 1 per cent. level; for enrolment, at the 5 per cent. level; and for interest and progressiveness, below the 5 per cent. level.

TABLE 2

MEAN E.Q.'S FOR UPPER AND LOWER QUARTERS OF DISTRIBUTIONS OF SCORES FOR SIX FACTORS, BASED UPON A SINGLE ATTAINMENT MEASURE COMPRISING BOTH COMPREHENSION AND ROTE SUBJECTS

Factor	Mean E.Q.		Difference
	Upper	Lower	
(iv) Intelligence	105.5	93.4	12.1
(iii) Socio-economic status	103.6	95.6	8.0
(v) Morale	104.1	96.1	8.0
(xiv) Enrolment	101.6	95.6	6.0
(x) Interest	102.0	98.8	3.2
(vii) Progressiveness	101.2	97.7	3.5

IV.—SUMMARY.

1.—This article has reported an investigation into the relationships of various characteristics of junior schools and their environments to one another and to the level of attainment in the final junior year, in a group of fifty mixed London junior schools.

2.—The schools were assessed as units for forty-two qualities by the techniques of measuring, testing, rating and questionnaire.

3.—The various aspects assessed included the school site and building, the homes and neighbourhood of the school, the Head, the Staff, the children, and the type of school organization.

4.—Five criteria of attainment were established, and eleven possible determinants of attainment. The inter-relationships among these sixteen factors were studied by the use of correlation and analysis of variance.

5.—The main factors determining level of attainment in the formal school subjects were found to be intelligence level, socio-economic status, large enrolment, and good morale.

6.—School morale itself was found to be determined not only by the attitudes and philosophy of the Head and Staff but by intelligence level and socio-economic status.

7.—Interest in school appeared to be a form of attainment and was found to be partly determined by intelligence level and socio-economic status. It tended to be stronger in voluntary schools.

8.—Adjustment or maturity of behaviour appeared to be closely related to morale and socio-economic status.

9.—Progressiveness was found to be little related to attainment in the formal school subjects, to morale, to adjustment, or to interest in school.

10.—Little relationship was found between class size and level of attainment.

V.—REFERENCES

- ALEXANDER, W. P. (1935): "Intelligence, Concrete and Abstract," *Brit. J. Psychol. Monogr. Suppl.*, 19.
- BURT, C. (1917): "The Distribution and Relations of Educational Abilities." (London: King).
- BURT, C. (1937): "The Backward Child." (London: University of London Press).
- BURT, C. (1939) "The Relations of Educational Abilities," *Brit. J. Educ. Psychol.*, 9, 45-71.
- BURT, C. (1947) "Mental and Scholastic Tests," Revised Edition (London: Staples).
- CAMPBELL, W. J. (1952) "The Influence of Home Environment on the Educational Progress of Selective Secondary School Children," *Brit. J. Educ. Psychol.*, 22, 89-100.
- CATTELL, R. B. (1948), "Concepts and Methods in the Measurement of Group Syntality," *Psychol. Rev.*, 55, 48-63.
- CATTELL, R. B. and WISPE, L. G. (1948), "Dimensions of Syntality in Small Groups," *J. Soc. Psychol.*, 28, 57-78.
- FLEMING, C. M. (1943), "Socio-economic Level and Test Performance," *Brit. J. Educ. Psychol.*, 13, 74-82.
- GARDNER, D. E. M. (1942), "Testing Results in the Infant School." (London: Methuen).
- GARDNER, D. E. M. (1950), "Long Term Results of Infant School Methods." (London: Methuen).
- GRAY, P. G., CORLETT, T. and JONES, P. (1951), "The Proportion of Jurors as an Index of the Economic Status of a District." (London: Central Office of Information).
- MANLEY, D. R. (1951): "An Experiment in the Standardisation of a General Information Test." (Unpublished thesis, University of London).
- SCHONELL, F. J. (1942): "Backwardness in the Basic Subjects." (London: Oliver and Boyd).
- THORNDIKE, R. L. (1941): "Observations of Behaviour of Children in Activity and Control Schools," *J. Exp. Educ.*, 10, 138-145 and 146-149.
- VERNON, P. E. (1950): "The Structure of Human Abilities." (London: Methuen.)
- WATTS, A. F. (1944): "The Language and Mental Development of Children" (London: Harrap.)

A SURVEY OF STUDENTS ADMITTED TO TRAIN AS TEACHERS IN SCOTLAND UNDER THE POST-WAR EMERGENCY SCHEME*

By JOHN SUTHERLAND
Moray House, University of Edinburgh

I.—Introduction. II.—The experiment. III.—Previous surveys. IV.—Results of the present survey: (a) the intelligence test scores; (b) success in professional training; (c) relationship between personality and teaching ability. V.—Summary. VI.—Acknowledgements. VII.—References.

I.—INTRODUCTION.

IN the years succeeding the 1939-45 World War, Scotland, like most other countries, was faced with a marked shortage of teachers. This was due mainly to the fact that the number of entrants to training declined during the war years but various other factors such as the raising of the school leaving age and the expected effort to reduce the size of classes were likely to make the shortage more acute.

In Scotland before the war, the position was that all teachers had to be trained and that all men, apart from teachers of practical subjects like Art, Music, Handwork, etc., had to be graduates. There are three main categories into which these students in training as teachers may be grouped:

- (1) First or Second Class Honours Graduates training to teach their special subjects in secondary schools. These are commonly referred to as Chapter V Students.
- (2) Students training for service in primary schools. These are commonly referred to as Chapter III Students. These consist of two groups.
 - (a) Students with Third Class Honours Degrees or with Ordinary Degrees. They are commonly referred to as Chapter III Graduates. They may also train to teach in the first three years of the secondary schools, subjects in which they are academically qualified.
 - (b) Non-graduate students (normally all women) who require an entrance qualification approximately the same as that required for University Entrance. They are commonly referred to as Chapter III Non-Graduates.
- (3) Students of the various practical subjects training to teach their particular subjects either in the primary school or the secondary school. They are commonly referred to as Chapter VI students.

The neatest method of referring to the various categories is by their Chapters (V, III or VI) and that is the method that will be used in this article supplemented, where suitable, by a further description of the particular group concerned. The number of students in each category is shown in Table I.

From October, 1943, the National Committee for the Training of Teachers, the body responsible for the training of all teachers in Scotland, operated a special Emergency Scheme designed to attract more recruits (mostly ex-Service men and women) to the teaching profession. Under this scheme modifications were made in the basic academic entrance qualifications in certain cases where a deficiency in normal qualifications was compensated by other qualities in the

* This survey was conducted on behalf of and with the co-operation of the National Committee for the Training of Teachers in Scotland.

applicant. Thus, suitable men without a university degree but with war service were admitted to train as teachers for service in primary schools. In addition, both men and women applicants for emergency courses who had approved national service were admitted, though their academic qualifications were not fully up to standard, *i.e.*, did not amount to a full Leaving Certificate or a full University Preliminary Examination pass. The Training Authorities also accepted the special war degrees awarded by the universities.

It was realized that many of the applicants would be unsuitable and, in order to select the suitable candidates, Provincial Selection Boards and a Central Selection Board were set up. Initially, all candidates were interviewed by one of these Boards or by their specially appointed interviewing officers. Later, those with a normal academic entrance qualification were exempted from the interview. The Boards then considered all the evidence and decided which candidates should be admitted. Selected candidates were awarded a maintenance grant which included remission of fees and was based on their special requirements.

Of the 357 Chapter III Non-Graduates admitted under the Emergency Scheme, 197 were men, of whom 144 did not have the normal academic entrance qualification for non-graduate women. Of the 160 non-graduate women admitted in this way, 46 did not have the normal academic entrance qualifications.

No special emergency training colleges were established. Those students admitted under the Emergency Scheme were absorbed into the seven training colleges already existing in Scotland. Thus, in the post-war years there was, in these training colleges, in addition to the normal entrants (largely women at this time), a great body of ex-Service men and women, some with normal academic qualifications and some without, some who would have trained as teachers in any case and some who would not, some with a normal academic background apart from the break for National Service and some who had left school at the age of fourteen and had since followed various callings before finally deciding to take up teaching.

This scheme was not without its critics. It was attacked on the ground that by reducing the academic entrance qualifications for admission to training it was lowering the status of the whole teaching profession in Scotland. Since the same protest against dilution of the teaching profession is being made with reference to another Special Recruitment Scheme now being operated, it is important that this point should be considered.

On the other hand, a very common criticism of teachers who enter the profession through the normal channels is that they are too academic in their outlook and that they have little or no knowledge of life. By this is generally meant that they have little knowledge of the likely conditions of employment and the social background of the majority of the children that they will teach. In the group of students admitted under the Emergency Scheme there were a number who had had this very type of experience that the others lacked. An opportunity was, therefore, provided of finding out at least to what extent this additional knowledge of life and this greater maturity affected the training colleges' estimates of their ability as teachers.

II.—THE EXPERIMENT.

It was, therefore, decided in 1948 that a comparison should be made between the students admitted under the Emergency Scheme and the normal entrants. An attempt was to be made to assess their intelligence, and certain

information about their ability as teachers and their success in academic and professional studies was to be supplied from the training colleges.

(a) *The Group Tested.*—All students actually in training as teachers in Scotland in the Spring of 1949 were tested. The numbers of students actually tested (excluding foreign students) are shown in Table I.

TABLE I
STUDENTS TESTED IN THE SURVEY

Training Category	Normal Entrants	Emergency Scheme Entrants	Total
Honours Graduates (Chapter V)	62	113	175
Ordinary Graduates (Chapter III)	270	258	528
Non-graduates (Chapter III)	1,670 ¹	357	2,027 ¹
Students of Practical Subjects (Chapter VI)	356 ²	296	652 ²
Total	2,358	1,024	3,382

¹ It should be noted that since the non-graduate course is a three-year course, this total is approximately three times the yearly output of non-graduate teachers at this time. The actual number of normal entrants in this category who completed their course in June, 1949, was 431.

² In the same way many of the Practical Subjects students take a two-year course. The number of normal entrants completing their course in June, 1949, was 202.

(b) *The Test Used.*—In view of the information available about it, Moray House Adult I. Intelligence Test was used. This is a typical verbal intelligence test of a level of difficulty suitable for adults (though perhaps not for very superior adults). It consists of 100 items each of which can be marked right or wrong. Most of the items are of the multiple choice type which are answered by underlining one of a series of choices. The others involve the minimum of writing, normally one letter or one digit, and occasionally one word. Two examples will illustrate the general types :

51.—Underline the word which means most nearly the opposite of **LIKELY** (possible, important, improbable, unimportant, reasonable, unfortunate).

87.—In a certain code A stands for 12, B for 6, C for 3, D for 2 and E for 1. Write the *letter* which represents the answer to this sum :

$$\begin{array}{r} A - C \\ \hline B + D + E \end{array}$$

The first standardisation of Moray House Adult I Test was carried out by Sir Godfrey Thomson in 1943.¹ He compared the scores obtained by 378 secondary school pupils age 15½–16½ with their intelligence quotients obtained on a standard Moray House test used at the age of 11+ about four years earlier. At this stage it was possible to obtain a properly representative sample of the population and the 11+ Intelligence Quotients, therefore, did measure the intelligence of these pupils compared with that of the whole population. On the basis of this comparison, Sir Godfrey Thomson derived the formula for Adult I.Q. :

$$\text{I.Q.} = \cdot 7 \text{ score} + 70.$$

It will be noticed that this formula makes no allowance for age. In other words, it implies no further growth in intelligence (or at least in the ability to answer verbal intelligence tests) after the age of $16\frac{1}{2}$ years. Some of the results obtained in this survey throw some doubt on this assumption and the formula must, therefore, be used with discretion.* Since the basis of comparison in most cases is raw scores, this particular difficulty will not arise.

During the past six years the test has also been applied to almost 600 university graduates along with the Otis Advanced Intelligence Test (another verbal test) and Ravens' 1938 Matrix Test (a non-verbal test). The results have been very consistent and reasonably high inter-correlations have been obtained. The correlation between the Moray House Test and the Otis Test is .78, between the Moray House Test and Ravens's Matrix Test .72 and between the Otis Test and Ravens' Matrix Test .66. Since both Moray House Adult I and the Otis Test are of a verbal nature, we would expect a higher correlation between them than with the Matrix test but the fact that the correlation with the Matrix test is as high as .72, seems to indicate that a very large proportion of what is tested by the test is common to the performance of both verbal and non-verbal intelligence tests. Lastly, correlations of over 0.85 have been obtained between scores obtained on Moray House Adult I Test at the ages of 15 and 16, and the scores of the same children on the Moray House 11+ intelligence test at that age.

It is reasonable to suppose, therefore, that the Moray House Adult I Intelligence test does, in fact measure to a large extent general intelligence and that it is a valid test of the intelligence of students in all categories.

(c) Other Information Supplied. After sitting the intelligence test, each student filled in a Questionnaire card which asked, in addition to his name, age, training college, university, etc., what occupations he had followed before entering upon training as a teacher and for what length of time, whether or not he possessed a Leaving Certificate or its equivalent and whether or not he would have become a teacher if there had not been a war. On the completion of training each training college supplied, for each student tested, a mark for practical teaching and a mark representing an overall estimate of ability as a teacher. These, and the score on the intelligence test, were also entered on the card.

For those students admitted under the Emergency Scheme who were interviewed before admission, a personality rating was also available and was entered on the cards by the training college authorities. For those accepted for training and therefore covered by the survey, only three main ratings, A, B and C, were awarded, but with pluses and minuses this became, in effect, an eight-point scale. An 'A' was awarded only to an outstanding student and a 'C' was a bare acceptance. Ratings of 'D' and 'E' were also awarded but candidates with such ratings were normally rejected.

This then provided the general background of each student in training at that time together with an estimate of his probable success as a teacher and it was possible to compare certain aspects of the two.

III.—PREVIOUS SURVEYS.

Two previous large-scale surveys of the intelligence of students in training as teachers in Scotland, have been carried out. In 1943, Moray House Adult I Intelligence Test was used¹ and all students then in training were tested while, in 1947, both Moray House Adult I Intelligence Test and Moray House Adult II Intelligence Test (a more difficult test) were used and only those students admitted under the Emergency Scheme were tested.

*A more recent standardisation of the test gives age allowances up to the age of 17 years 6 months.

In the following tables of results, the figures from the 1943 and the 1947 surveys are included for the sake of comparison.

IV.—THE RESULTS OF THE PRESENT SURVEY.

(a) *The Intelligence Test Scores.*

(1) Comparison of Emergency Scheme Entrants with Normal Entrants.—In Table II are shown the Median, Upper and Lower Quartile scores obtained by each category of students in each survey. An approximate conversion table is also included. The most striking feature of these results is the overlap between categories. There were, for example, students from every category with scores higher than that of the middle honours graduate and one honours graduate had a score considerably below that obtained by over 75 per cent. of the whole group (a score giving an I.Q. lower than is now usually necessary to gain admission to an academic secondary school!). Taking the groups as a whole, however, we find that the honours graduates score highest in the test. They are followed by the ordinary graduates, the Chapter III non-graduates and the Chapter VI students, in that order.

TABLE II

COMPARISON OF SCORES OBTAINED IN MORAY HOUSE ADULT I INTELLIGENCE TEST BY NORMAL AND EMERGENCY ENTRANTS TO THE TEACHER TRAINING COLLEGES IN SCOTLAND.

	Category	Number	Intelligence Test Score		
			Upper Quartile	Median	Lower Quartile
Honours Graduates	1943 Normal Entrants	40	88	80	67.5
	1947 Emergency Scheme Entrants	101	91.1	81.8	71.1
	1949 Normal Entrants	62	90.3	82.8	71.7
	1949 Emergency Scheme Entrants	113	88.6	81.2	72.2
Ordinary Graduates	1943 Normal Entrants	222	83	73	62.5
	1947 Emergency Scheme Entrants	202	85.8	74.1	62.9
	1949 Normal Entrants	268	85.6	76.2	65.4
	1949 Emergency Scheme Entrants	258	85.0	73.9	60.9
Chap. III Non-Graduates	1943 Normal Entrants	739	73	63	53
	1947 Emergency Scheme Entrants	612	78.0	67.6	56.1
	1949 Normal Entrants	1670	75.2	66.0	56.8
	1949 Emergency Scheme Entrants	357	81.2	69.8	60.1
Students of Practical Subjects (Chap. VI)	1943 Normal Entrants	432	69.5	58.5	50.1
	1947 Emergency Scheme Entrants	161	66.8	57.2	44.6
	1949 Normal Entrants	356	71.4	60.3	51.2
	1949 Emergency Scheme Entrants	296	70.6	59.9	50.9

CONVERSION TABLE FOR PERSONS AGED 17 YEARS 6 MONTHS UPWARDS (BASED ON THE RECENT STANDARDISATION).

Score	30	40	50	60	70	80	90
I.Q.	93	100	106	112	120	130	141

TABLE III

SCORES OBTAINED ON MORAY HOUSE ADULT I INTELLIGENCE TEST BY SEPARATE CATEGORIES OF STUDENTS OF PRACTICAL SUBJECTS (CHAPTER VI).

	Category	Number	Intelligence Test Score		
			Upper Quartile	Median	Lower Quartile
Handwork	1947 Emergency Scheme Entrants	78	65.7	56.4	42.0
	1949 Normal Entrants	29	73.7	60.5	53.9
	1949 Emergency Scheme Entrants	145	70.8	58.7	49.6
Art	1947 Emergency Scheme Entrants	12	62.8	54.5	39.5
	1949 Normal Entrants	46	65.6	55.2	46.3
	1949 Emergency Scheme Entrants	39	68.6	58.9	48.3
Music	1947 Emergency Scheme Entrants	15	65.1	56.4	38.9
	1949 Normal Entrants	16	81.2	77.0	69.5
	1949 Emergency Scheme Entrants	13	68.9	58.3	51.6
Commercial Subjects	1949 Normal Entrants	26	80.1	66.2	55.3
	1949 Emergency Scheme Entrants	17	79.1	68.7	60.1
Physical Education	1943 Normal Entrants	110	74	60	52
	1947 Emergency Scheme Entrants	56	68.9	60.1	50.1
	1949 Normal Entrants	177	71.0	61.0	52.4
	1949 Emergency Scheme Entrants	77	69.9	60.7	52.1
Domestic Science	1943 Normal Entrants	322	69	59	50
	1949 Normal Entrants	61	60.3	54.3	48.3
	1949 Emergency Scheme Entrants	5	63.3	58.3	55.1

Table III shows the Median, Upper and Lower Quartile scores obtained by the separate categories of Chapter VI. students. A study of this table reveals that those students training as teachers of commercial subjects (many of whom are, of course, graduates) score highest and those taking domestic science score lowest.

The Emergency Scheme students when compared with the normal intake, hold their own in all categories. In one category, the non-graduates, the Emergency Scheme students, are superior to the normal intake, the difference being statistically significant.

Inspection of the scores of the upper quartiles in each distribution revealed that the emergency entrants provided just as many outstanding students as were contributed by the normal entrants. In the non-graduates the balance is again decidedly in favour of the emergency entrants.

An examination of the lower quartiles showed the Emergency Scheme students again holding their own with the possible exception of the Ordinary Graduates where the Emergency Scheme group contributed a rather larger number of less able students. In the Chapter III non-graduates on the other hand, the balance was again in favour of the Emergency Scheme students. Among the Chapter VI students there was little difference, except in the 1947 Emergency Scheme group, which had a rather larger number of poorish students but few very poor students.

The Emergency Scheme non-graduate group, who are superior to the normal intake, are of particular interest since they supply for the most part the real additions to the teaching profession. Most (though not all) of the graduates being

trained under the Emergency Scheme would have come to the training colleges in any case, but this scheme was the only way open to the non-graduate men who formed the bulk of this group. It attracted to the profession many men (and women) from established positions in life who had now decided that teaching was their vocation but who, because of their commitments, could not have trained without financial assistance and who, because of their lack of academic qualifications, would not have been admitted under normal regulations.

In view of the special interest attached to the group of non-graduate men, a further comparison was made of their scores with those obtained by other groups (all 1949) :

	Median
Chapter III non-graduate men (all Emergency Scheme).....	68.2
Whole Emergency Scheme Chapter III non-graduate group (men and women)	69.8
Normal Chapter III non-graduate group (all women)	66.0
Graduate men (normal and Emergency Scheme entrants)	75.4

The differences between the non graduate men and the graduate men, and between the non-graduate men and the normal non-graduates are significant. Thus, the Chapter III non-graduate men admitted under the Emergency Scheme more than hold their own with other non-graduates, though they do not measure up to the standard of the graduates. (The best group of the Chapter III non-graduate students consisted of Emergency Scheme women).

Two other comparisons were made for the whole of the group admitted under the Emergency Scheme (men and women). In certain cases students with a sub-normal academic entrance qualification were admitted if there were compensating factors. In this category there were 144 men and 46 women out of 357 Chapter III non-graduates admitted under the scheme. The scores of these groups were compared with the scores of the normal non-graduate entry. Their median scores were men 66.0 and women 64.5. The median for the normal entrants was 66.0 while the median for all students admitted under the Emergency Scheme was 69.8. Thus, while those students with a sub-normal admission qualification compare unfavourably with the Emergency Scheme group, as a whole, the men hold their own with the normal entrants and the scores of women are not significantly different.

TABLE IV

COMPARISON OF SCORES OBTAINED ON MORAY HOUSE ADULT I INTELLIGENCE TEST BY THOSE STUDENTS WHO WOULD NOT HAVE BECOME TEACHERS HAD IT NOT BEEN FOR THE WAR AND THE EMERGENCY SCHEME WITH THOSE OBTAINED BY NORMAL ENTRANTS.

Category		N.	Median
Graduates	Men	63	77.0
	Women	23	72.0
	All	86	76.7
	Normal Entrants....	330	77.5
Non-Graduates	Men	197	68.8
	Women	29	71.4
	All	226	69.0
	Normal Entrants....	1670	66.0
Chapter VI ..	Men	107	62.4
	Women	10	57.8
	All	117	62.0
	Normal Entrants....	356	60.3

On the questionnaire which all students filled in they were asked whether or not they would have become teachers had the war not taken place. It was known that many people admitted under the Emergency Scheme, though quite properly qualified for admission under the scheme, would have become teachers in any case and it was desired to find out something about those students who had entered upon training as a result of the war and the Emergency Scheme. The mean scores obtained by these students are shown in Table IV.

Among the graduates none of the differences is significant. Among the Chapter III non-graduates, while none of the differences is significant, it should be noted that both the men and the women who entered teaching as a result of the Emergency Scheme, were superior to the normal entrants. Among the Chapter VI students the differences are again extremely small.

(2) Comparison of Emergency Scheme entrants drawn from Different Occupations.

An attempt was next made to determine which occupation or class of occupations, appeared to contribute the most satisfactory recruits to the teaching profession, and which contributed the largest number. Occupations were classified according to the schedule shown in Table V, which is a very slightly modified form of that used in the recent Scottish Mental Survey and by the Royal Commission on Population.² In cases where a student had followed more than one occupation he was classified according to the occupation he had followed for the longest period. For the purpose of this analysis, the students were divided into three categories, graduates, non-graduates and students of practical subjects (Chapter VI) and only occupations which had been followed by ten or more students were included in the analysis. Army service was counted as an occupation only if no other occupation had been followed. Any occupation followed for less than six months was not included in the analysis.

TABLE V
CLASSIFICATION OF PREVIOUS CIVILIAN OCCUPATION.

1. Professional Class (including School Teachers).
2. Employers employing 10 or more workers.
3. Self employed or employing less than 10 workers.
4. Non-manual workers paid by monthly salary or less frequently.
5. Manual workers paid by monthly salaries or less frequently.
6. Clerks paid by weekly wage.
7. Non-Manual workers paid weekly.
8. Manual workers in skilled occupations who are paid weekly.
9. Manual workers in semi-skilled occupations who are paid weekly.
10. Unskilled manual workers or labourers paid weekly.
11. Manual workers skill unknown.
12. Farmers.
13. Agricultural workers.
14. Home Duties.
15. Student.

The results of this analysis are shown in Table VI. Among the graduates, those coming from clerical occupations have the highest intelligence, while those drawn from skilled manual occupations (trades, etc.) have the lowest, though it is not significantly lower than that of the normal entrants. Those with Army service only, who form the bulk of the group, score almost exactly the same as the normal entrants. Of those who had followed another occupation other than Army service before entering training, the vast majority were drawn from professional and clerical occupations.

TABLE VI

MEDIAN SCORES OBTAINED ON MORAY HOUSE ADULT I INTELLIGENCE TEST BY EMERGENCY SCHEME STUDENTS DRAWN FROM VARIOUS CLASSES OF OCCUPATIONS.

Category	Previous Occupation	Number	Median
Graduates	Professional Class ..	41	77.8
	Clerical	64	79.5
	Skilled Manual	19	75.3
	Army Service Only	233	77.3
Non-Graduates (Chap. III)	Professional Class	31	71.3
	Clerical	147	70.0
	Skilled Manual	58	64.5
	Monthly Salary Non-Manual ..	13	68.3
	Paid Weekly Non-Manual ..	22	69.5
	Semi-skilled Manual	18	67.0
Students of Practical Subjects (Chapter VI) ..	Army Service Only	47	67.9
	Clerical	31	62.0
	Skilled Manual	149	57.9
	Army Service Only	83	61.5

Among the non-graduates the clerical occupations again yield students well above the average as do also the professional occupations, with those coming from skilled manual occupations again the lowest. Those who had previously followed a clerical occupation form by far the largest class.

In the Chapter VI category, the clerical group of occupations again provides the students with the highest scores though they are far below the corresponding group in the other categories, and the skilled manual group of occupations again provides the students with the lowest scores. It must be remembered, however, that such a test may be unsuited to this particular type of student. The skilled manual workers, as one would expect, form the largest group.

(b) *Degree of Success in Professional Training.*

Comparison of Emergency Scheme entrants drawn from various occupations with normal entrants.

In the course of a student's training, various assessments of his ability are made. The one which means most to him and which is probably the most important is the assessment of his practical skill as a teacher. This is essentially a subjective estimate on the part of the student's supervisor of teaching practice. Various cross checks are made within each college to standardise the marking of various individuals and samples of all students from all colleges are heard each year by Her Majesty's Inspectors of Schools. In this way it is possible to approximate to a reasonably uniform standard of marking. To obtain a high mark a student must have a sound grasp of his subject and must know how to handle children.

Final teaching marks for all students in training in the Spring of 1949 were submitted as soon as the students had completed their training. All but a very few had done so by June, 1951. The marks obtained by students admitted under the Emergency Scheme were then analysed by previous occupation and were compared with those obtained by normal entrants. The results of this analysis are shown in Table VII.

Perhaps the most marked feature of all the distributions is the very large number of students who obtained marks between 65 and 69, and the very small value of the interquartile range. It will also be noted that there are no students with marks below 55.*

The small range of marks is probably the result of a subjective attempt to assess an ability which involves so many diverse factors. There are certain basic qualities which a teacher must possess, but beyond that there are very many ways in which a person may be a good teacher. In addition, the teaching supervisors are attempting to assess the student's potential aptitude and it is inevitable that when prediction is involved a certain amount of conservatism should enter into the assessment.

The cutting short of the distributions at 55 is due to the fact that a student with a teaching mark below this value is unlikely to be certificated by the training colleges. It is true that some students do not attain this standard. Such a student may have his course terminated without the award of a certificate or he may himself realise that teaching is not his vocation and decide to seek other employment. On the other hand, there may be many reasons for a student discontinuing his training, so it was decided to include in the distributions only those who completed the courses. The number of students thus omitted is very small.

Owing to the limited range of marks, it is difficult to select any one occupation previously followed by a reasonable number of students which clearly has contributed better students than other occupations. Among the graduates, however, the professional and clerical classes again produce students above the average while those who previously followed occupations in the skilled manual group are again below the average. Among the Chapter III non-graduates the numbers in each group are on the whole so small that little can be deduced. In the groups comprising a reasonable number of students, the clerical group is above average and the skilled manual group is also slightly above average. Among the Chapter VI students all groups admitted under the Emergency Scheme are above the average of the normal entrants.

In the last line of Table VII are shown the teaching marks obtained by the non-graduates with sub-normal entrance qualifications. If we compare these marks with those obtained by the other non-graduates, we find that this group is statistically superior to the normal entrants and that there is no significant difference between their scores and these of the emergency entrants as a whole. It is evident, therefore, that whatever qualities of personality or background were possessed by these students and which, in the view of the interviewing boards, compensated for their lack of academic attainment, have also been approved of by their teaching supervisors.

Pinsent³ and Turnbull⁴ in researches carried out in 1933 and 1934, and Saer⁵ in a similar research carried out in 1941, found that pre-college teaching experience produced little or no difference in the teaching mark obtained at college and Pinsent, in the same research, found that within the limits of age of his group, there was no connection between teaching mark and age. The case of the present group of ex-service students, however, is rather different. The feature common to them all was that, having tried one or more occupations they had now decided that teaching was the profession they wished to follow and they were determined to make a success of it. Many of them were men and women of mature experience and judgment and of well-developed personality.

*Percentage marks do not appear on the student's certificate. They are used as a guide in the award of letter marks on a seven-point scale.

TABLE VII

TRAINING COLLEGE ASSESSMENT OF TEACHING ABILITY OF STUDENTS DRAWN FROM VARIOUS OCCUPATIONS.

Category	Previous Occupation	Number	Teaching Mark		
			Upper Q	Median	Lower Q
Graduates	Professional Class	41	72.4	69.0	65.3
	Clerical	61	72.3	68.4	66.1
	Skilled Manual	18	68.2	66.4	64.7
	Army Service Only	233	71.0	67.7	65.2
	Normal Entrants	336	70.4	67.5	65.1
Non-Graduates	Professional Class	24	73.0	69.5	66.0
	Clerical	138	72.9	68.6	65.7
	Skilled Manual	57	71.9	68.3	65.8
	Monthly Salary Non-Manual	13	73.0	69.9	66.4
	Paid Weekly Non-Manual	22	72.9	68.4	65.4
	Semi-skilled Manual	16	71.5	67.5	63.5
	Army Service Only	46	70.3	67.3	64.8
	Normal Entrants	1529	71.3	67.9	65.3
Students of Practical Subs. (Chap. VI)	Clerical	18	72.7	69.5	65.8
	Skilled Manual	131	73.2	68.7	66.3
	Army Service Only	67	73.0	68.8	65.6
	Normal Entrants	356	70.2	66.4	62.7
Non-Graduates with Sub-Normal Entrance Qualifications		190	74.0	69.0	66.1

Just because of this, it may well be that some of them will not develop any further. With the young student coming straight from school or university, one has to make allowance for his potential development. It is necessary to estimate the extent to which he likely to develop, and this, in some cases, may be considerable. Training college assessments are based on the performance of comparatively immature students and allowance must inevitably be made for future development. Thus, the training college scale of assessment may not be a valid scale upon which to measure the ability of more mature students.

Hence, although these older more mature ex-service students may be better performers at the present time, it may well be that the younger normal entrant will reach an equally high standard once he has had time to mature. Nevertheless, qualify the conclusions as we may, the fact remains abundantly clear that in all categories of students the Emergency Scheme has contributed to the teaching profession of Scotland a number of men and women with marked teaching ability.

It was thought that, having discontinued their studies for some time, some of the students admitted under the Emergency Scheme might find it difficult to hold their own in the more academic subjects of the course. A comparison in this field was very difficult since in very few cases did Emergency Scheme students and normal entrants in any large number follow the same course. An exception was found in the Edinburgh Training Centre where both Emergency Scheme and Normal Chapter III graduate entrants followed the same courses in Education and Psychology. The results of this comparison are shown in Table VIII.

TABLE VIII

COMPARISON OF FINAL MARKS OBTAINED IN EDUCATION AND PSYCHOLOGY BY NORMAL AND EMERGENCY SCHEME CHAPTER III GRADUATE STUDENTS.

Category	Education		Psychology	
	N.	Median	N.	Median
Emergency Scheme Entrants	49	64.2	30	64.5
Normal Entrants	62	66.8	30	69.5

Though neither difference is statistically significant, the slightly lower scores obtained by the Emergency Scheme entrants in *both* subjects would seem to suggest that they do experience some difficulty in re-establishing habits of study.

(c) *Relationship between Personality and Teaching Ability.*

Many studies have attempted to find a method of predicting teaching ability other than by assessing the student's ability in the classroom, but with little success. In this survey, teaching ability as assessed by the training college staffs, was correlated with intelligence, with education and with psychology, but in all cases the correlations were not statistically significant.

For the students who were admitted under the Emergency Scheme, and who were interviewed before admission, there was available a subjective personality assessment made in most cases by the interviewing officers and finally awarded by the Selection Boards. These assessments, made on an eight-point scale, were correlated with the final teaching marks. The results are shown in Table IX.

TABLE IX
CORRELATION BETWEEN PERSONALITY AND TEACHING ABILITY

Training College	No. of Students	Correlation
Glasgow	163	.403
Edinburgh	120	.310
Aberdeen	61	.308
Dundee	96	.412

These coefficients are not high but they are considerably higher than those obtained in most other researches. They must, however, be qualified in two respects. Each student was normally interviewed by three officers, two of whom usually had training college experience, and one of whom was a member of the staff of the Scottish Education Department. A certain 'halo effect' might, therefore, exist between the two subjective assessments since the interviewers would tend to be looking for the same qualities which they had looked for or were looking for in teachers in training. Secondly, the estimate of teaching ability used as criterion was the final teaching mark awarded by the training colleges and it is not known as yet how this correlates with future success as a teacher.

Tudhope,⁶ in 1942, found a correlation of .81 between training college estimates and an assessment by His Majesty's Inspectors at least three years later. Later researches, however, have failed to confirm his figures.

None the less, the correlation coefficients in Table IX are high enough to warrant serious consideration and to suggest that it may become possible to predict reasonably well a student's probable future success.

V.—SUMMARY.

In the Spring of 1949, all students then in training in Scotland (numbering 3,382) sat an intelligence test and filled in a questionnaire. This included students of all categories, honours and ordinary graduates, non-graduates and teachers of practical subjects. These students have since been followed up to the completion of their training courses.

Some of these students were admitted under what was known as the Emergency Scheme, some were normal entrants. Some Emergency Scheme students were admitted with a sub-normal academic entrance qualifications if the interviewing boards considered that there were adequate compensating factors, some had normal entrance qualifications; and some became teachers only because of the special facilities offered by the Emergency Scheme while some would have become teachers in any case. Lastly, those admitted under the Emergency Scheme had followed a variety of occupations apart from war service before entering upon training.

The following are the principal findings of the survey:

1.—Considerable overlap of intelligence was found between the different categories of students, but on the whole the honours graduates were of highest intelligence, followed by the ordinary graduates, the non-graduates (Chapter III) and the teachers of practical subjects, in that order.

2.—In each training category the students admitted under the Emergency Scheme held their own with the normal entrants.

3.—Of the Emergency Scheme students, those who were admitted with a sub-normal academic entrance qualification compared unfavourably in intelligence with the Emergency Group as a whole, but held their own with the normal entrants.

4.—Those who came into training only because of the inducements offered by the Emergency Scheme did as well as normal entrants.

5.—Of all the occupations previously followed by the Emergency Scheme students the clerical occupations yielded the most intelligent students, while the skilled manual occupations yielded those with the lowest intelligence.

6.—In teaching ability as assessed by the training colleges, the Emergency Scheme students held their own with the normal entrants. Those non-graduates admitted with a sub-normal academic qualification did at least as well as the whole group of Emergency Scheme non-graduates and better than the normal non-graduate entrants.

7.—The clerical occupations yielded the best teachers and the skilled manual occupations the poorest. The differences, however, were not large.

8.—In their professional studies of Education and Psychology, the Emergency Scheme graduate students did not do so well as the normal entrants, but the differences were not large enough to be statistically significant.

9.—The personality assessment awarded to certain Emergency Scheme students as a result of a personal interview, showed a significant agreement with the training college assessments of their teaching skill.

VI.—ACKNOWLEDGEMENTS.

Thanks are due to the National Committee for the Training of Teachers in Scotland on whose behalf this survey was conducted, to their Executive Officer and to the Directors and Principals of the Training Colleges for their willing co-operation and assistance at all stages of the survey.

VII.—REFERENCES.

- ¹ THOMSON, G. H. : "The Distribution of Intelligence among University and College Students," *Brit. J. of Educ. Psych.*, Vol. XV, Part II (1945).
- ² THE SCOTTISH COUNCIL FOR RESEARCH IN EDUCATION : *The Trend of Scottish Intelligence* (Univ. of Lond. Press, 1949.).
- ³ PINSENT, ARTHUR : "Pre-College Teaching Experience and other Factors in the Teaching Success of University Students," *Brit. J. of Educ. Psych.*, Vol. III, Parts II and III (1933).
- ⁴ TURNBULL, G. H. : "The Influence of Previous Teaching Experience on Results obtained by Students in a University Department of Education," *Brit. J. Educ. Psych.*, Vol. IV, Part I (1934).
- ⁵ SAER HYWELA : "Further Investigation of Pre-College Teaching Experience and other Factors in the Teaching Success of University Students," *Brit. J. of Educ. Psych.*, Vol. XI, Pt. 3 (1941).
- ⁶ TUDHOPE, WILLIAM B. : "A Study of the Training College Final Teaching Mark as a Criterion of Future Success in the Teaching Profession," *Brit. J. of Educ. Psych.*, Vol. XII, Pt. 3 and Vol. XIII Part I. (1942).

THE USE OF AN INTEREST TEST IN 11 PLUS SELECTION

By STEPHEN WISEMAN

Senior Lecturer in Education, University of Manchester.

I.—Introduction. II.—Subjects and tests. III.—Scoring methods for interest tests. IV.—Factor analysis. V.—A technique of selection. VI.—Summary. VII.—Acknowledgements. VIII.—Appendix.

I.—INTRODUCTION.

IN a previous paper¹ the construction and partial validation of an interest test was described. The test exists in two versions, one for boys and one for girls, and each consists of ninety-six statements, divided into 'sets' of six, to each of which the subject responds 'like,' 'don't know,' 'dislike,' or 'not done.' Within each set of six, the subject then marks the one liked best, and the one liked second best. From these latter responses, two scores are obtained, a practical score, P, and an academic score, A. Correlations between P-score and A-score were found to be -0.66 (boys) and -0.36 (girls) on separate samples of over 200 children of 13 plus, who were candidates for technical school selection. Reliabilities for the separate scores ranged from 0.88 to 0.91.

It seems likely that the chief value of a test such as this will be to help in selection at 11 plus, and in particular to assist in discriminating between pupils more likely to succeed in secondary technical schools and those better fitted for grammar schools. The results reported for children of 13 plus are, therefore, interesting but not convincing. It is obvious that further results are needed from children of 11 plus. Such an analysis has now been carried out, and the results are reported below.

II.—SUBJECTS AND TESTS.

A sample of 200 boys from ten Salford primary schools was tested in 1953. All the boys were in the final year of primary education, and were taking the grammar school selection examination, part of which had already been held before this additional testing. The full range of ability was not represented, as head teachers were asked to omit the lowest fifth in the ability range. The restriction of range is evident from the mean quotients and standard deviations given in Table I. The sample was given the Jenkins scale of Non-Verbal ability, the Moray House Space Test I, the National Foundation for Educational Research Spatial Test I, Peel and Lambert's General Information Test I, and the Devon Interest Test. In addition, results were available from the 11 plus selection examination (Moray House tests of Intelligence, English and Arithmetic). Although an essay is used in Salford for grammar school selection, it is given to only part of the age-group, and it was not included in the analysis as some children in the sample had no essay mark.

III.—SCORING METHODS FOR INTEREST TESTS.

The Peel and Lambert interest test uses a single score which assesses a "bias of interest."² This is a difference-score showing the bias of practical

¹ FITZPATRICK, T. F., and WISEMAN, S.: "An Interest Test for Use in Selection for Technical Education," this *Journal*, 1954, 24, pp. 99-105.

² LAMBERT, C. M.: "Symposium on Selection of Pupils for Different Types of Secondary School," this *Journal*, 1949, 16, p. 69.

interests over academic interests. It will be remembered that in this test the subject is presented with sets of six questions, out of each of which he must do three. Three questions are practical in content, and three academic. His P-score is the total number of practical questions he gets right; similarly for A-score and academic questions. Since each of these scores will correlate highly with general scholastic ability, this factor is eliminated by using the formula

$$\frac{P - A}{P + A} \times 100.$$

TABLE I

RANGE OF ABILITY IN THE SAMPLE.

Test	Mean Quotient	S.D.
MH Test 52	107.7	10.830
MH Space Test I	103.2	11.190
National Foundation Spatial Test I	102.7	12.690
Jenkins' Non-Verbal Test	106.9	12.575

It should be noted that in the Peel and Lambert test the subject can choose *only* practical or academic items. In the Devon test, each set of six items contains, in addition to two practical and two academic items, one 'strong distractor' (e.g., 'playing football') and one 'social' item (e.g., 'being a member of a scout unit'). The use of a difference score on this test would produce some curious results. For example, two children both have a difference score of 3; one arrives at this by achieving a P-score of 25 and an A-score of 22, while the other has a P-score of 3 and a zero A-score. These children have most dissimilar interest patterns. The first one has only chosen *one* social item or strong distractor out of thirty-two possible; the second has chosen *all* such available items with the exception of three (or possibly two) practical items. In other words, the first (on this evidence only) seems to be a good bet for either grammar or technical school, whereas the second should be recommended for neither. A trial analysis of 13 plus results showed that the existence of such dissimilar patterns was by no means rare.

It seems clear that the results from the Devon test can best be used by expressing them as two separate scores, a P-score and an A-score. These raw scores may easily be converted into standard scores having a mean of 100 and a standard deviation of 15 and can thus be used in conjunction with the quotients from standardised tests. In Section V of this paper a way in which such scores may be used in selection is suggested.

IV.—FACTOR ANALYSIS.

The crucial problem in allocating children to technical schools and grammar schools is to find some measure which will differentiate the abilities and aptitudes required for these two types of courses. The relative lack of success of spatial tests in solving this problem has been due to the fact that such tests appear to measure primarily intelligence (*g*) and only secondarily spatial ability (*k*). In other words, the superiority of a particular child in this test may be due to a superiority in general intelligence (required for *both* types of school) rather than a superiority in spatial ability (required for the technical school). The *g*-element in the test score is useless for differentiation, and merely duplicates the information we already have from the straightforward intelligence test. What is required ideally is a measure (or measures) which can differentiate

efficiently, without being contaminated with intelligence or attainment factors which are adequately covered by other tests in the battery. The degree of independence of test-scores can most easily be demonstrated by the use of factor analysis. Since the present battery contains both verbal and non-verbal intelligence tests, two space tests, two interest tests and two attainment tests, a factor analysis ought to provide important information about the inter-relationships of the various abilities and aptitudes.

Two factor analyses were, in fact, carried out since it was desired to investigate the separate P and A scores given by the Peel and Lambert test as well as the difference score. The inter-correlations between the various measures are given in Table II. It will be noticed that low, or negative, correlations are found between the Devon P and A scores and the Peel and Lambert difference score on the one hand, and the tests of intelligence, spatial ability and attainment on the other. Thus it seems as if the interest tests are relatively independent of the other tests, as had been hoped. This is borne out by the factor analysis (details of which are given in the Appendix).

TABLE II
INTERCORRELATIONS. N=175 (Decimal points omitted).

	2	3	4	5	6	7	8	9	10	11
1—MH Intelligence	774	748	739	712	487	-001	013	656	526	-138
2—MH Arithmetic		647	580	586	408	-019	-055	229	348	-101
3—MH English			568	576	299	-036	139	256	510	-176
4—Jenkins' Non-Vb.				718	587	015	-020	341	394	-021
5—MH Space					700	129	-017	375	368	-010
6—NF Space						144	-129	430	238	126
7—Devon P							-550	-087	089	126
8—Devon A								-149	177	-216
9—Peel P									032	
10—Peel A										
11—Peel Difference										

NOTE.—Of the sample of 200, only 175 completed all the tests.

The first factor, or general factor, represents an over-all average, as it were, of the common elements of all the tests. It is, therefore, very largely an 'intelligence-attainment' factor (or *g : ed* in Vernon's notation). The presence of the interest tests—and to some degree the space tests—prevent it from being a clear *g : ed* factor. We can, however, by rotating axes, 'purify' it by ensuring that the intelligence and attainment tests are given maximum loadings on this factor.

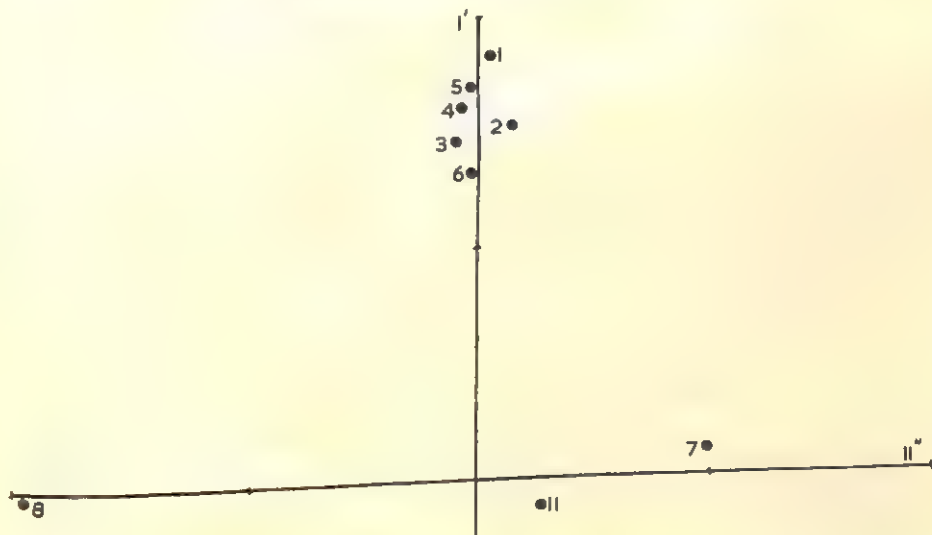
The second factor to emerge is the first 'differentiating' factor, showing positive loadings on some tests and negative on others. Such bipolar factors are, of course, just what we are looking for for technical selection, since they may provide just that element of 'bias' which we require. In other words, the general factor indicates tests which will help us to differentiate between children destined for technical and grammar schools on the one hand, and modern schools on the other, while the bipolar factors will separate the technical pupils from the grammar pupils. It is, therefore, of considerable interest to note that the first, and most influential, of these bipolar factors is quite clearly an *interest* factor, while the second bipolar (factor III) is a space factor. Rotation of axes was again made to produce comparable factor patterns which would most clearly

bring out the relationships between the tests. This was done very simply by ensuring that the loadings of the space tests on the space factor were maximized.

What may we conclude from this factor analysis? The first important piece of information will be found in the last line of each of the tables in the Appendix, which gives 'percentage variance.' It will be seen that the interest factor (II) has a value nearly twice as large as that for the space factor (III). This means that the differentiating factor of interest is, in this battery, almost twice as influential as the space factor. We may hope, therefore, that discrimination between technical and grammar school will be improved considerably by using an interest test in addition to a space test.

DIAGRAM

PLOT OF FACTORS I' and II", USING PEEL AND LAMBERT DIFFERENCE SCORE
(Tests numbered as in Table II)



Secondly, it appears as if the two separate scores (P and A) of the Devon test are likely to be more efficient in selection than the single difference score of the Peel test. The diagram shows the loadings on factors I and II after rotation of the Peel test. The vertical distance of each point from the horizontal axis shows the loading of each test on the general (*g:ed*) factor. The intelligence, space and attainment tests are all highly saturated with this factor, while the three interest scores are independent of it. By using interest tests, then, we are not duplicating information already available: interest scores are not contaminated with the intellectual-educational factor. Distances measured horizontally from the vertical axis show the loadings on practical interests (measured to the right) and on academic interests (measured to the left). None of the tests 1-6 has other than a negligible loading on this factor, whereas the Devon P and A scores have large opposite loadings. The Peel difference score has a disappointingly low practical loading—although it should be noticed (see the Appendix table) that it has quite a substantial loading on the space factor. We may conclude, therefore, that the Devon P and A scores, when added to a battery already containing a space test, add new and independent information.

V.—A TECHNIQUE OF SELECTION.

Selection at 11 plus within the tripartite system of secondary education, shows some similarity to personnel selection in the Forces. For example, when selecting various types of aircrew (pilots, navigators, etc.), a battery of tests may be given to the total entry, and different combinations of tests, with different weights may be used to arrive at 'pilot-scores,' 'navigator-scores,' etc. What tests should contribute to each score, and what weights to be used, can be decided by applying multiple regression technique to follow-up data. In exactly the same way we could arrive at a 'technical-score' and a 'grammar-score' for each child of 11 plus. Until adequate follow-up results were available for a particular authority, it would be necessary to base our 'test-combinations' on what we know from research-in-general on predictive-power, factorial-content, etc. Differential weighting is probably, at this stage, unrealistic.

Looking at the problem from this point of view, it seems likely that scores on the basic tests of verbal intelligence, English and Arithmetic should contribute to *both* technical and grammar-scores, since it is generally agreed that both types of secondary school demand a reasonably high level of intelligence and attainment. We certainly do not wish to support the pre-war notion that the technical school should be a second-best. Research on K-factor suggests that the space-test score should form part of the technical-score, and one would also add P-score from the interest test. What little evidence we have¹ at present on the predictive value of the English essay, suggests that this should form part of the grammar-scores. To this would be added A-score from the interest test. Thus we arrive at a selection battery of six tests, yielding a technical-score and a grammar-score for each child, and each composed of five elements, as shown below:

<i>Technical-Score</i>		<i>Grammar-Score</i>	
Intelligence	}	Basic Tests	{ Intelligence
English			{ English
Arithmetic			{ Arithmetic
<hr/>			<hr/>
Space-test			Essay
Practical interest			Academic interest

The essay and interest scores would, of course, be standardised with a mean of 100 and S.D. of 15, so as to be comparable with the quotients from the 'basic tests.'

In order to see how such a scheme would work out in practice, it was applied to the Salford sample of 175 boys. It was assumed that about 25 per cent. of children were to be allocated to grammar or technical in roughly equal numbers. Since the sample was restricted in range, it seemed likely that a total intake of 30 per cent. from it would roughly reproduce these conditions. A border-zone was fixed, centred so as to give 30 per cent. intake, and extending over four points of average quotient. It should be noted that no essay marks were available for the sample, so that this was missing from the grammar score. Thus, the grammar score was the aggregate of four quotients, and the technical score of five. The border-zone width for grammar school was, therefore, 16 points, and for technical 20 points. Table III shows the distribution of the sample.

¹ See WISEMAN, S., 1952: "Problems Connected with the Selection of Children for Secondary Schools," Ph.D. thesis, Manchester University, for evidence that the essay score significantly increases prediction of School Certificate English Literature, but does not do so for the examination as a whole. To use the essay only for 'grammar-score' does not, of course, imply that ability to write connected prose is not required for technical education. The high correlation between composition and objective English tests means that the objective test can reject those children whose composition ability is very weak.

TABLE III
SCORE DISTRIBUTION OF 175 BOYS

		Technical Score		
		Low	High	
Grammar Score	High	560	580	
	472	3 C	6 B	16 A
	456	10 F	5 E	4 D
Low		109 K	14 H	8 G

(Border-zone enclosed in double rules)

Selection might now proceed as follows :

Area A—16 boys. These are clear admissions to *either* grammar school or technical school. Decision will be on (a) parents' choice, (b) school recommendation.

Area B—6 cases. Clear admissions to grammar school.

Area D—4 cases. Clear admissions to technical school.

If, in either area, the parents' choice is opposed to the allocation, the child takes his chance in the borderline (E).

Area C—3 cases. Clear admission, grammar.

Area G—8 cases. Clear admission, technical.

If parents' choice opposed in either area, modern is the only alternative.

Area E—5 cases.
Area F—10 cases
Area H—14 cases

} The border-zone, scrutiny of school-reports, and any other additional machinery in use—e.g., interview and observed activities (Northumberland, Devon). Roughly *half* these cases will be recommended to the selective schools. Movement only to technical from H, only to grammar from F.

Area K—109 cases.—Clear admission to modern.

It will be seen that a fair degree of discrimination has been achieved by the test scores, and if an essay mark were used in addition, discrimination is likely to increase. The areas A and E contain those children who cannot be allocated without additional evidence: a total of $16 + \frac{5}{2}$ out of a total intake of 51 or 52 boys. Thus, roughly 63 per cent. of the intake has been classified by the test battery.

VI.—SUMMARY.

1.—The Devon Interest Test has been applied to a sample of 175 boys of 11 plus, together with tests of intelligence, attainment and spatial ability, and Peel and Lambert's interest test.

2.—Factor analysis shows that the interest factor is, in this battery, considerably more influential than the space factor.

3.—The interest tests are shown to be independent of the general factor of intelligence-attainment.

4.—The Devon practical and academic scores appear to be more effective for differentiating between technical and grammar pupils than does the Peel and Lambert difference score, particularly if a space test is included in the selection battery.

5.—A method of technical-grammar allocation is suggested which appears to result in positive classification of about 60 per cent. of pupils.

VII.—ACKNOWLEDGEMENTS.

Grateful acknowledgement is made to the Director of Education, Salford, and to the members of the Consultative Committee on Technical Selection, under whose auspices the enquiry was made. Particular thanks are due to Mr. A. B. Peel, Inspector of Schools, who made all the arrangements for the testing of the boys, and to the heads of the schools concerned for their ready co-operation.

VIII.—APPENDIX.

CENTROID ANALYSIS

A.—ANALYSIS USING THE DIFFERENCE-SCORE FOR THE PEEL AND LAMBERT TEST.

After four iterations, the estimated and obtained communalities were in close agreement. Rotation was done graphically. Factors I and II were plotted and the axis rotated 21° clockwise until the axis of the general factor ran through the cluster of tests 1-6. Factors II' and III were then plotted and a new III' axis was run through Test 6 (N.F. Space Test). This needed a rotation of 26° anti-clockwise. The original and rotated loadings are shown in the table below :

	Original Loadings			Rotated Loadings		
	I	II	III	I'	II'	III'
1—MH Intelligence	·888	·245	·236	·917	·023	·251
2—MH Arithmetic	·728	·237	·245	·765	·071	·238
3—MH English	·757	·077	·292	·735	—·051	·349
4—Jenkins N.V.	·757	·295	—·105	·813	—·042	—·096
5—MH Space	·773	·373	—·194	·856	—·021	—·205
6—NF Space	·550	·447	—·472	·673	—·009	—·520
7—Devon P	—·148	·532	·016	·053	·501	—·227
8—Devon A	·329	—·916	—·213	—·021	—·968	·236
11—Peel Difference	—·156	·224	—·226	—·066	·139	—·319
Percentage Variance . .				42·5	13·5	8·6

B.—ANALYSIS USING SEPARATE P AND A SCORES FOR PEEL AND LAMBERT.

This analysis was a little troublesome, owing to the communality for Test 8 rising to well over 1·0. After nine iterations, however, a reasonably good fit was obtained, with only Test 8 showing a difference of over 0·01 between estimated and obtained communalities. The same principles of rotation were adopted. For I and II the rotation required was $14^\circ 30'$ clockwise, and for II' and III 37° anti-clockwise. The original and rotated loadings are shown below :

	Original Loadings			Rotated Loadings		
	I	II	III	I'	II'	III'
1—MH Intelligence	·962	·170	·190	·974	·053	·198
2—MH Arithmetic	·703	·114	·274	·710	·218	·179
3—MH English	·768	—·100	·358	·719	—·016	·460
4—Jenkins N.V.	·773	·204	·000	·799	·002	—·002
5—MH Space	·783	·294	—·069	·832	·029	—·109
6—NF Space	·609	·449	—·358	·702	·011	—·456
7—Devon P	—·098	·479	·136	·025	·473	—·185
8—Devon A	·243	—·860	—·423	·020	—·969	·200
9—Peel P	·409	·380	—·037	·491	·190	—·190
10—Peel A	·516	—·174	·204	·456	—·114	·342
Percentage Variance . .				42·4	12·6	7·3

In each analysis the third factor was significant using the criterion of Burt and Banks.

GROUP RESEARCH WITH POST-GRADUATE STUDENT LABOUR IN A DEPARTMENT OF EDUCATION

By E. W. HUGHES

Organizer of Research, Institute of Education, Leicester.

I.—*Introduction.* II.—*Planning and method of working.* III.—*Results and comments on procedure.* IV.—*Summary.* V.—*Appendix.*

I.—INTRODUCTION.

During the session 1952/3, a large-scale group research project on the Relationship of School to Industry and Commerce was conducted in the Department of Education of the University College, Leicester, in collaboration with the Institute of Education. The students taking part were all graduates training as teachers in the Department of Education.

The project was intended :

- (i) To test the possibility of using Post-Graduate Student Labour in large-scale research.
- (ii) To limit and co-ordinate the number of individual demands on the schools made by a large body of students in search of material for their thesis topics.
- (iii) To give some training in research methods.
- (iv) To show that research work and course work could be usefully integrated—in this case on the problems of Vocational Guidance.

The group project itself was designed to follow up the research work of Gustav Jahoda¹ and Cora Tenen² in connection with the adjustment of the adolescent to industry and commerce, with a view to estimating :

- (a) the extent of the adolescent's knowledge of industrial conditions and opportunities.
- b) The effect of environmental pressures on the school leavers in determining their choice of occupation.

This article will describe how the work was organized so that the students could be assessed both for their group participation and their individual contributions to the project, and thereby provide assessors with an alternative approach to the extended essay requirement for the Diploma (Certificate) in Education.

The products of the group research itself will be reported on in another paper.

Throughout this research it was clearly recognized that in no way could we regard the students as Research Assistants and that whilst we were providing a framework in which the student could give proof of his abilities, we still had an obligation to teach.

Furthermore, it was held that a basic principle of such group work should be individual specialization within the main framework, with a full appreciation, by the students, of the place of individual topics in the whole and a knowledge of the contribution of individual findings to the whole.

¹ GUSTAV JAHODA : " Adolescent Attitudes to Starting Work, *Occup. Psych.*, June, 1949.

² CORA TENEN : " The Adolescent in the Factory," *Brit. J. of Educ. Psych.*, June, 1947.

II.—PLANNING AND METHOD OF WORKING.

At the beginning of the Autumn Term, 1952, a draft research scheme on the Relationship of School to Industry and Commerce, was submitted to the officials of the local education authority of a large town, and to a group of twenty-two diploma students, as a basis for discussion.

Briefly, this proposed that all the secondary modern school leavers for 1953 in the town should be asked to write an imaginative essay, "My First Day at Work" (very much as in Jahoda's research, mentioned previously) and complete a rather full and searching questionnaire. The essay would be examined for evidence of knowledge of industrial conditions and opportunities, and the questionnaire would be designed to throw light on the factors affecting choice of occupation.

The scheme called for the co-operation of the heads and staffs of the thirty-one secondary modern schools in the setting of the essay and questionnaire to the 2,000 or so leavers (boys and girls), and anticipated consultation with the local education authority secondary school organizers, the youth employment officer and the town planning research officer, as the investigations impinged on local education authority policies and made demands on other survey materials relating to industrial opportunities.

Initial Stages.

One afternoon per week of the Autumn term was spent in discussing the draft material and in the shaping of the final form of the research. During the first two weeks the earlier work of Gustav Jahoda and Cora Tenen was described and the general setting of the problem of transference and adjustment as it affected the youth, the school, the parents, the youth employment service, and industry and commerce, was explained and discussed. This was much in the nature of coursework. As the students began to face the practical problems of the research, it was soon obvious that some specialization would be necessary and they agreed to divide into sub-groups as under :

Sub-Group (a)—Background research and published material.

Sub-Group (b)—Questionnaire design and analysis.

Sub-Group (c)—Essay analysis proposals.

Sub-Group (d)—Map work.

Sub-Group (e)—Related research on motivation.

This grouping was voluntary and not directed.

When the various reactions of the local education authority officials to the original proposals became available, they were submitted to the students as a whole and then referred to the appropriate sub-group for action. For example, the grounds for objections to the form of certain items in the questionnaire were made known to the complete group and the questionnaire sub-group took the necessary actions to meet the criticism of the local education authority officials, bearing in mind the general expression of opinion in the group as a whole.

Thus, all the members of the main group knew what was happening and what to expect from their own and each other sub-group. The exception was Sub-Group (e) which was introduced to cover the details of a parallel research, to which the students could be switched in event of failure to obtain the necessary co-operation from the schools.

A number of essays were available from an earlier investigation and these were used by Sub-Group (c) in framing their proposals for essay analysis.

Progress.

By the end of November, 1952, agreement had been reached on the detailed form of the questionnaire and of the essay instructions, but because of

possible overlapping with the youth employment officers conducting school interviews at this time, the local education authority suggested that work should start in the schools in the Spring term.

At this stage the students reviewed their work to date in the form of written reports relating to the various sub-groups and by Christmas, the members of the group were in a position to decide whether or not to proceed with this research project for the special paper requirement of their diploma.

The blank questionnaire, instructions to schools, and essay analysis sheets were duplicated in the Institute of Education and early in 1953, the organizer of research contacted the heads of the schools directly, requesting their support.

The response was extremely encouraging; all the thirty-one schools co-operated and over 2,000 sets of scripts were produced, with occasional delays. By mid-February all the students participating had received complete sets of questionnaires and essays for analysis purposes.

Four meetings were arranged to discuss detailed methods of analysis of questionnaires and essays and to pass on any recommendations of the Sub-Groups (b) and (c) on questionnaire and essay analysis. It was during these sessions that advice on simple research techniques was offered. At the request of the students these recommendations on the main points to examine, agreed policies, and useful techniques, were duplicated for future reference. It is important to notice that these were recommendations and not instructions, as the student was always free to proceed as he wished.

Generally, each student aimed at a three-stage report:

- (a) Analysis of essays from the school.
- (b) Analysis of questionnaires.
- (c) Personal aspect of the work, particularly interesting the student and possibly reflecting sub-group specialization.

Reporting Stage.

Early in the Summer term the policies for the shaping of the separate theses and the submission of the project as a whole to the External Examiner were decided by discussion and recommendation. The members of the various sub-groups were asked to meet together and prepare a brief summary of their specialist activities and findings. These reports were then read to the main body of students as a whole and critical comments invited. This provided opportunities for each member of the main group to be brought up to date with the work of other members. On the basis of the discussion which ensued, slight adjustments were made to the original summaries and these were then deemed fit to submit to the external examiner. The summaries were used as introductions to the work of each sub-group and preceded the actual theses of the group workers in the mass of material submitted to the external examiner.

Because of the peculiar difficulties associated with the reproduction of maps, duplicated copies of maps were issued to all group members for small additions relating to the individual schools. The originals were based partly on the work done by the map group, supplemented by information from the Town Planning Department and the Youth Employment Service. The duplication was the responsibility of the Institute of Education.

The research organizer read all the theses on completion, and prepared a comprehensive statement which showed for each student:

- (a) Volume of material (questionnaires and essays) dealt with.
- (b) The extent and estimated quality of general participation.
 - (i) In respect of essays.
 - (ii) In respect of questionnaires.

- (c) The nature and quality of personal (including specialist) contribution.
- (d) Comments on unusual or outstanding features.
- (e) Final rating.

External Examiner.

The External Examiner received :

- (a) A general description of the research project and its conduct, together with copies of original documents, i.e., draft material, final material and all instructions.
- (b) The collected theses arranged in order of the sub-groups and preceded by the prepared summaries (referred to above) for each sub-group.
- (c) The research organizer's comprehensive reports on the work of the students.

After examining this material the external examiner decided that the group project called for group review. He critically examined the idea of the project with the members and was challenging in his stimulation to discussion. Generally, the project was evaluated as an introduction to research methods, as an educative process, and as a useful aid to local education authority and other officials, by way of the products of the research.

Submission to Local Education Authority.

Early in July it was possible to submit the students' theses to the local education authority organizers and youth employment officer. Thus, reports on twenty of the thirty-one schools were available by the end of the session and were submitted to the heads of the schools early in the following session.

Collation.

The students realized that the next valuable stage of the research project would be the collation of the data from the thirty-one schools. They recognized that the time factor would prevent them from completing this last phase of the research and that it would fall to the Institute of Education to devise methods for producing the final documents. The action taken in this respect will be noted later in this article.

The submission of the students' theses to the local education authority, as reports on the research findings for individual schools, represented a natural end to the first phase of this research.

III.—RESULTS AND COMMENTS ON PROCEDURE.

General.

During the early stages of this research the organizer faced a dual problem in connection with his obligations to the students and to the local education authority. Firstly, having persuaded a group of students to consider the project, could he be sure that the local education authority and schools would support him and also respond early enough to provide the raw material for the students' extended essays? Secondly, having obtained the support of the local education authority and schools could he now be sure that a sufficient number of students would agree to continue with the project after the first term of planning?

In this particular case, the contact with the schools was delayed slightly at the request of the Youth Employment Service to meet an administrative difficulty, but enough material was available by mid-February to accommodate all the students and the material which was delayed from a few schools did not interfere with the students' participation.

(In this connection it should be noted that Sub-Group (e) was outlining another related research which could have been switched to accommodate many of the students had the original project failed to materialize.)

As far as the student participation was concerned, one resigned after the second week of the project, and two others, although interested in the problem, did not want to do more than join in the planning and discussion. One other student left the department in the Spring term. Of the eighteen students reporting on twenty schools, two did not use their reports for thesis purposes, but did the work as an extra activity in the Department.

The Institute of Education was able to interest five practising teachers from its Further Diploma Groups in this project and this group, together with the research staff of the Institute, took on the task of dealing with the scripts from the schools outstanding.

As a first stage to collation, it was decided to make statistical abstracts for all schools and for this reason careful checks have been made of the students' original theses and their rough work tables. This enables us to make a number of realistic comments by way of evaluation of these theses.

Students' Theses.

In the first place the theses included opinion. This was sometimes, but not often, exaggerated. Only one school protested about such expressions of opinion, the others tended to ignore it and concentrate on the factual aspects of the reports, interpreting the material in light of their unique knowledge of the conditions existing. It can be seen, therefore, that from the point of view of the schools, as distinct from students and assessor, the statistical abstracts prepared for collation purposes will prove useful.

The re-checking of the factual material showed that the accuracy of the students' statistical work was high. The main difference between the original and revised tables were often due to matters of opinion in classification of activities as distinct from false entries. Most of the discrepancies would disappear if greater direction of students' activities were to be considered appropriate. This again would make for greater uniformity in the calculation of inter-column agreements. Occasional tables had to be recast for collation purposes when the student diverged significantly from the general method of classification, but on the whole the broader classifications and groupings were not seriously affected.

Generally, the discrepancies in the statistical work were sufficiently small to enable the general conclusions and trends in the uncorrected reports to be accepted in their own right.

Another feature which the collation work has brought to light is the actual weight of work involved in the abstraction of the agreed data. There was a considerable range in the number of sets of scripts from the different schools (25-104). In two cases students dealt with two schools each, but there was still a wide range in volume of work. 60-70 sets of scripts (both essays and questionnaires) was considered to be a good allocation of work, but several sets reached 100 (i.e., 100 questionnaires, 100 essays). In one or two cases it is felt that the sheer weight of work resulted in the omission of some of the inter-question correlations, whilst in other cases the light load gave opportunities for considerable cross checking and critical examination of data. However, one thing is certain, there is no evidence that any student had an easy task and, in fact, much evidence to the contrary.

The research organizer's assessment (see entries under heading Reporting Stage, earlier) made allowance for volume of work and other participations, but the difficulty of combining volume of work with quality and kind of specialization, destroys the possibility of reaching any degree of precision in the final

ratings. The writer, therefore, prefers the verbal comments under the separate headings, in his report, to the final ratings for the theses as a whole.

The reading of the students' special contributions, i.e., third sections of their theses, made it abundantly clear that the project offered a wide range of special features. Amongst the specializations adopted were the following themes :

Youth employment officer/careers' master dilemma.

Motivation problems in (i) industry ; (ii) teacher selection.

Secondary modern school policy in relation to this adjustment problem.

Agricultural bias in town schools and rural schools contrasted.

Free and controlled essays in this research.

The design and effectiveness of questionnaires.

Methods of representing proximity of industry in diagrammatic forms.

And the like. In providing opportunities for this kind of study and comment within a larger project and in a real life situation, the project has succeeded. It may, however, be criticized from the point of view that with the exceptions of the members of the sub-groups themselves, the rest of the students had substitute experience only, in the sense that they listened to the reports of the work of others, with occasional discussions. Their direct experience was limited to their own field of specialization. This might be a particularly heavy objection in relation to the work of the Sub-Group (a)—Background research and published material, although here it must be pointed out that the research organizer tried to remedy some of these deficiencies in the introductory work of the first two weeks. Furthermore, there was evidence, in the form of bibliographies, of wider reading by the students in relation to the topic of their own specialization.

It is true to say that the students were shown in a practical way the ramifications of a large-scale research and the sort of problems to be faced in a further degree research study. They all did a constructive piece of work by taking on the analysis of a self-contained part of the main body of material and in specializing in a feature of the research. They achieved at least as much by these methods as they would have done in the same time, on non-project work. There is no doubt either, that the opportunities for discussion did bring to light points of view and ideas which otherwise might never have occurred to both students and organizer.

There was evidence of general satisfaction with the project on the part of the members of the group, especially in relation to the content of the scheme and the research techniques.

The practical nature of the work brought with it a sense of urgency. The students are now aware of the size of some of the research problems in the field of vocational guidance. They are aware of the difficulties of balancing the policies of public bodies in designing an investigation. They are aware of the drudgery entailed in research work and (I hope) some of the elation. They are aware of the problems of large-scale reproduction of material and the need for reporting results to co-operating bodies. Above all, they are aware of the dangers of easy generalization and the difficulties of getting a straightforward answer to what appears to be a simple question.

It remains for the final report on the collated material to round off this satisfaction with evidence of useful research findings.

These findings will appear in general report form for the interested officials and organizations concerned and in the form of statistical abstracts for the separate schools. In this way the research section of the Institute of Education will feel that it has met its two obligations to the students and to the local authority.

IV.—SUMMARY.

1.—This article has described how a body of post-graduate students in a Department of Education joined in a large research project, on the Relationship of School to Industry and Commerce, involving thirty-one schools and over 2,000 adolescent boys and girls.

2.—It shows how it was possible to give scope for individual specialization and initiative within the framework of the project and so provide a means for assessing the individual student's contribution, both to the main theme and to the topic of his specialization.

3.—The article shows that the initial planning and later stages of analysis made it possible to integrate course work of a background nature with a training in research methods, whilst still pursuing a valuable research theme on the vocational aspects of adolescent adjustment.

4.—Not least amongst the lessons of the project were those which concerned the research worker's obligations to the co-operating bodies.

5.—Rigorous direction of the investigation was not considered appropriate and recommendation was the key to our attitude.

6.—The methods adopted represented an alternative approach to the special paper requirement for the Diploma (Certificate) in Education.

ACKNOWLEDGEMENTS.

I should like to thank all those people who co-operated in the original research—Relationship of School to Industry and Commerce—which has made this article possible.

V.—APPENDIX.

GENERAL TIME TABLE

Stage	Organizer	Group of Students	L.E.A. Official or School	
PLANNING AND PREPARATION.	Prepares draft of scheme and submits to Students and L.E.A.			
	Background to research explained to students. Adjustment problem discussed. Meets sub-groups and co-ordinates findings. Duplicates Questionnaires and Instructions.	Consider draft proposals and divide into sub-groups for specialization. Modify original drafts in light of L.E.A. recommendations and group discussions. Proceed with Sub-Group work. Christmas reports and decision on full participation in scheme.	Consider draft proposals and recommend adjustment in light of own policies. Agree final draft.	Autumn Term. One afternoon per week.
TRIALS AND ANALYSES.	Contacts Heads of Schools for support.			
	Issues scripts to Students in rotation and discusses method of work.	Commence analysis. Attend four meetings on methods	Schools complete Essays and Questionnaires.	Spring Term.
REPORTING AND ASSESSMENT.	Discusses form of presentation of theses for Dept. and External Examiner	Complete theses in three stages. Sub-groups prepare summaries for submission to group and assessor.		Summer Term.
	Assesses theses. Presents project to External Examiner	External Examiner discusses project with students.		
			Theses to L.E.A. Theses to Schools.	
	Material (including theses) to Institute of Education for collation and Statistical Abstracts.			Session 1953/54.

THE RELATIVE EFFICIENCY OF INTELLIGENCE AND ATTAINMENT TESTS AS PREDICTORS OF SUCCESS IN GRAMMAR SCHOOLS

By JACK WRIGLEY

Lecturer in Education, Queen's University, Belfast.

I.—Introduction. II.—Survey of previous research. III.—The Lancashire follow-up: (a) description of experiment; (b) results; (c) discussion of results. IV.—The Northern Ireland follow-up: (a) description of experiment; (b) results; (c) discussion of results. V.—Conclusions. VI.—Summary.

I.—INTRODUCTION.

A RECENT article by B. F. Hobby¹ has questioned the generally accepted view of educational psychologists, that the intelligence test is superior both to English and Arithmetic tests, as a predictor of success in grammar schools. A good case is made out by Hobby, and there seems little doubt that some misleading statements have been made in the past. For example, Hobby has questioned a statement made by Professor Schonell² in an important review of researches. Hobby has also pointed out that less evidence on this question has been published than might be imagined from the confident assertions of some psychologists. Nevertheless, it is possible for individual statements to be proved wrong, or shown to be misleading, and for the central thesis still to be valid. This article attempts to contribute an unbiased survey of previous research and then presents new evidence from two separate follow-up experiments.

II.—A SURVEY OF PREVIOUS RESEARCH.

One difficulty encountered in an attempt to survey previous research, is to decide which experiments to leave out on grounds of inadequacy or unsuitability. For example, Hobby throws doubt on some researches which have used teachers' estimates to measure success in the grammar school; he implies that it is more scientific to use marks in a public examination such as School Certificate. But this is not necessarily true. It should be recognised that any criterion of success in school is likely to have some limitations. Most researchers are well aware of this fact; whether they seek to use teachers' opinions or try to evolve a criterion from marks obtained in a public examination, they are merely using the best estimates available. It is not necessarily more scientific to use marks rather than estimates on such a matter as success in a school. If teachers' estimates of success are used, what is important is the way in which the estimation is done, and it is usually impossible to find this out from the published report of an experiment. Similarly, it might be that, in a particular selection procedure, one of the tests used was a rather inadequate example of its kind (for example, a poor intelligence test or a badly constructed arithmetic test). Yet it would probably be unscientific to exclude such results from a comprehensive survey, on these grounds. The principle adopted, in this survey, has

¹ HOBBY, B. F.: "Intelligence or English?—A Consideration of Prognostic Values in the 11+ Transfer Examination," *Higher Educ. Journ.*, 1953, V, IX, N. 3, p. 9.

² In this *Journal*, 1949, V, XIX, Pt I, p. 24, SCHONELL wrote: "Thus Amos, Collier, Hughes, Godfrey Thomson, Stubbins, showed that an intelligence test is the best single prognostic element in an examination test battery."

been to use all results which could possibly be adapted to our purpose, and to assume that differences in the ability of examiners to set suitable papers, and differences in the suitability of estimates of success in the school, would be random, operating in favour of neither intelligence tests, nor arithmetic nor English papers.

Two of the earlier researches^{1,2} referred to by Professor Schonell in the statement questioned by Hobby, have been excluded from this survey. This is because the intelligence test results were not used in the actual selection examination. It has been shown by F. Sandon³ that it is not valid to compare a follow-up correlation obtained from a test which was not used for actual selection purposes with one that was so used. There is a tendency for the test which was not used to yield a higher correlation than it should, just because it had a non-selective function. In all the researches used in this survey the three tests, English, Arithmetic, and Intelligence, were each used in the actual selection of candidates, though in some cases the intelligence test results were used for border-line pupils only.

It is well-known that the absolute size of a correlation coefficient varies with the selective nature of the sample; higher values usually being obtained with heterogeneous groups than with more homogeneous ones. W. G. Emmett⁴ has described a method of correcting for selectivity and has used it in a number of follow-up studies. The method is particularly suited to enable comparisons to be made between follow-up experiments in different areas, but it is not necessary to use it for measuring the relative predictive value of the component parts of selection examinations. In any case many experimenters have not obtained the data required to make corrections for selectivity, so that unless some valuable data is to be discarded uncorrected correlations must be used. In this survey uncorrected correlations have been used in every case.

In two experiments it has been found necessary to adapt the results of the researches for our purpose. In the first of these, T. E. Stubbins' reported correlations between Arithmetic, English, and Intelligence tests, comprising an 11+ examination on the one hand, and nine subjects in School Certificate on the other hand. The nine subjects were English Language, English Literature, History, Geography, German, French, Mathematics, Physics, and Chemistry. Taken as a whole, these nine subjects appear to give a valid measure of general success in the school. Consequently, the weighted average of the correlations with each of Arithmetic, English and Intelligence tests, was taken as a measure of the predictive value of these three parts of the 11+ examination. The average was calculated by transforming r to Fisher's Z , taking a weighted average of Z , and then transforming back to r . The average values so obtained were:

	Arithmetic	English	I.Q.	N (average)
Group I,174	.232	.202	154
Group II133	.265	.243	148

¹ AMOS, A. D.: "Examination and Intelligence Test Forecasts of School Achievement," *Brit. J. Educ. Psych.*, 1931, V. I, p. 73.

² COLLIER, J. W.: "The Predictive Value of Intelligence Tests for Secondary Education," *Brit. J. Educ. Psych.*, 1933, V. III, p. 65.

³ SANDON, F.: "Selection by a Nearly Perfect Examination," *Annals of Eugenics*, 1936.

⁴ EMMETT, W. G.: *An Inquiry into the Prediction of Secondary School Success* (London Univ. Press, 1942).

⁵ STUBBINS, T. E.: "The Prognostic Values of the Subjects of a Secondary School Entrance Examination," *Brit. J. Educ. Psych.*, 1940, X, p. 16.

It would be wise to treat the figures with caution for any method of averaging correlations has limitations. Nevertheless, if the figures obtained are interpreted in a relative way, and regarded merely as a convenient way of summarising complex results, they will perform a useful function.

A similar process has been carried out with results obtained by E. A. Peel and D. Rutter.¹ In their research, they used two sets of School Certificate marks : (1) English Language, English Literature, French and Mathematics ; (2) Geography, Mathematics, Biology, and Physics, as criteria of success in the school. The first group appears to be a rather better set of subjects from which to derive a broad estimate of success in the grammar school.

The average correlations, calculated as before, were :

Arithmetic	English	I.Q.	No. in group
.118	.392	.417	234

In the other experiments, reported by Emmett,² and Emmett and Wilmot,³ a satisfactory estimate of general success in the grammar school has been used and their results are taken as they stand. Success in the grammar school was estimated by teachers in the experiment reported by Emmett, and by School Certificate results (marks in English Language, English Literature, History, French and Mathematics) in the Emmett-Wilmot study. Table I shows the complete results obtained in the researches reviewed in this article.

TABLE I
CORRELATIONS WITH SUCCESS IN GRAMMAR SCHOOL.

	Arithmetic	English	Intelligence	Number of Children
Stubbins (1)174	.232	.202	154
Stubbins (2)133	.265	.243	148
Emmett330	.335	.454	765
Emmett118	.392	.417	234
Peel-Rutter424	.476	.499	128
Emmett-Wilmot (1)408	.488	.553	153
Emmett-Wilmot (2)				1582
Average284	.355	.422	

The average correlations, calculated as before, and given at the foot of Table I, express the whole of the available evidence in a simple way. It is clear that the order of merit for predictive efficiency is (1) Intelligence test ($r=0.422$), (2) English ($r=.355$), (3) Arithmetic ($r=0.284$). These average correlations are, in effect, calculated from the results of 1582, children and should be reasonably stable. Thus, the generally accepted view of success in the grammar school, is intelligence test is the best predictor of success in the grammar school, is confirmed by this survey. Further comment on the figures obtained is reserved until the results of two follow-up experiments carried out by the author have been added.

¹ PEEL, E. A., and RUTTER, D. : " The Predictive Value of the Entrance Examination as Judged by the School Certificate Examination," *Brit. J. Educ. Psych.*, 1951, XXI, p. 30.

² EMMETT, W. G., *op. cit.*

³ EMMETT, W. G., and WILMOT, F. S. : " The Prediction of School Certificate Performance in Specific Subjects," *Brit. J. Educ. Psych.*, 1952, XXII, p. 52.

The results obtained in W. McClelland's¹ well-known experiment in Scotland, have been excluded from the calculations because the correlations obtained were not exactly comparable. The Scottish experiment yielded a group of 461 children in the senior secondary school and the correlations obtained were: Arithmetic .671, English .699, Intelligence .700. These correlations are higher than any others summarised in this article. This is because (a) selection for senior secondary schools was not very stringent at the time of the experiment and consequently the follow-up groups were more variable in ability than the groups in England; (b) a correction was made for restriction of range of ability in the second and third year groups.

III.—THE LANCASHIRE EXPERIMENT.

(a) *Description of experiment.*

The sample studied consisted of 320 children forming the whole of three consecutive years in a co-educational grammar school in Lancashire. The performance of these children in the entrance examination, comprising a verbal and non-verbal intelligence test, and objective English and Arithmetic tests, was compared with their later success in the school. The children were considered as three groups, corresponding to the year in which they took the entrance examination (Group I in 1948, Group II in 1947, Group III in 1946), and for these groups the inter-correlations between all tests and success in the school were calculated. The relative effectiveness of the separate parts of the entrance examination, as predictors, was measured by their correlations with the criteria of success.

Two criteria of success were used in this experiment. In 1949, when the present writer was a member of staff of the school in question, success in the school was measured by a combination of marks in terminal examinations and teachers' estimates. This criterion was thus established after one, two and three years in the school, for the three groups concerned. The teachers were asked to make general estimates of the pupils' progress on a five-point scale. Three independent estimates were made for each pupil and an average taken where opinions differed. Full details of the methods used in constructing this measure of success and of the whole experiment at this stage can be found in an unpublished thesis.²

The follow-up was continued until all three groups of children had taken the General Certificate of Education at the Ordinary level (in 1951, 1952, and 1953). By this time the numbers of children left in the three groups were: Group I 90, Group II 70, Group III 67. The second criterion of success was obtained by taking an average mark for all subjects taken at the Ordinary level, any child taking less than four subjects being excluded from the experiment. The General Certificate of Education is less suitable than the former School Certificate for evolving a broad criterion of success in the school. The criterion which was evolved was again merely the best available in the circumstances.

(b) *Results.*

(1) The correlations of the component parts of the entrance examination with both criteria of success are given in Table II.

¹ McCLELLAND, W. : *Selection for Secondary Education* (Lond. Univ. Press, 1942).

² WRIGLEY, J. : "The Prediction of Success in the Grammar School," unpublished M.Ed. thesis, Manchester University, 1951.

TABLE II
CORRELATIONS WITH SUCCESS

	First Criterion (Estimates in School)				Second Criterion Marks in G.C.E.			
	Group I	Group II	Group III	Aver- age	Group I	Group II	Group III	Aver- age
Non-Verbal Intelligence ..	.152	.135	.093	.131	.226	.035	.291	.188
Arithmetic335	.179	.243	.261	.335	.338	.037	.254
English279	.030	.198	.178	.377	.040	.158	.191
Verbal Intelligence294	-.062	.197	.155	.204	.181	.254	.211

(2) The correlations between the two criteria of success were: Group I .684, Group II .667, Group III .776.

(c) *Discussion of results.*

(1) Whichever estimate of success is taken, the Arithmetic test is the best predictor of success, on the average. This finding appears to be a rare one; none of the experiments surveyed previously have produced such a result.

(2) There was little difference between English and verbal intelligence as predictors of success.

(3) The correlations are all rather small, probably due to the highly selective nature of the sample. For Group II the correlations are so small as to indicate almost a breakdown in prediction. Such evidence as there was, including personal knowledge of the children concerned, suggested that the breakdown in prediction was due more to the peculiar nature of the sample, rather than to defects in the entrance examination tests.

(4) The most suitable results to be added to the previous survey of follow-up studies would appear to be those obtained using marks in the General Certificate of Education as a criterion of success. If non-verbal intelligence is omitted, the figures to be added to the established results in Table I are:

	Arithmetic	English	Intelligence	No. in sample
	.254	.191	.211	227
Correlation with success				

IV.—THE NORTHERN IRELAND EXPERIMENT.

(a) *Description of experiment.*

In this experiment the sample studied comprised all the County Antrim children (N=472) who sat the 1948 Qualifying Examination (the name given to the 11+ examination in Northern Ireland), and entered grammar schools. The Qualifying Examination is set by the Ministry of Education in Northern Ireland and almost all intending grammar school entrants sit the examination. In 1951 the 472 children sat Junior Certificate and 320 of them went on to take Senior Certificate in 1953. These two examinations are also central examinations administered by the Ministry of Education. Junior Certificate is taken after either three or four years in the grammar school and Senior Certificate two years later. Junior Certificate is a similar type of examination to the former School Certificate in England, but is of a lower standard. The standard of Senior Certificate is somewhere between the former School Certificate and Higher School Certificate. It has some of the attributes of both these examinations for instance, a candidate has to take more subjects than he would in the English

Higher School Certificate, but he may take some of them at Advanced level. Both examinations are capable of providing a suitable criterion of general success in the school.

The criterion used in Junior Certificate was an average mark calculated from the following eight subjects: (1) English, (2) Latin, (3) French, (4) History, (5) Geography, (6) Arithmetic, (7) Algebra and Geometry combined, (8) Experimental Science. Before averaging, the marks in each subject were scaled to the same mean and standard deviation so that each subject carried the same weight in the final average. This criterion is broad and unbiased, and seems to be an accurate measure of success in the grammar school.

Senior Certificate is not quite so suitable a medium for providing a valid estimate of general grammar school success. In the procedure adopted, the marks in the various subjects were first adjusted so as to make the average mark for all subjects the same for all Northern Ireland entrants. Then an average mark was calculated using every subject taken by the candidates in Senior Certificate. This procedure has certain disadvantages. The assumption that the average performance in all papers can be represented by the same numerical mark is almost certainly not fully justified. Further, the criterion includes all subjects taken by the candidates and some of these may not be as valid measures of academic success as others. Nevertheless, after tentative experiments with other possible criteria, it seemed that the one outlined above suffered from the least defects.

In contrast to the Lancashire experiment the Arithmetic and English papers set in the Qualifying Examination were of the 'traditional' type. The Arithmetic paper was of a high standard of difficulty and contained a number of quite difficult problems. Results for English Language and English Composition are given separately. The intelligence test used was a Moray House one, and was used in the actual selection procedure for border-line candidates only.

(b) *Results.*

Table III shows the correlations obtained between the component parts of the Qualifying Examination and success in Junior and Senior Certificates. The correlation between success in Junior Certificate and success in Senior Certificate was 0.74.

TABLE III
CORRELATIONS WITH SUCCESS IN THE GRAMMAR SCHOOL.

	Junior Certificate N=472	Senior Certificate N=320
Arithmetic543	.38
English Language489	.34
English Composition.....	.427	.25
Intelligence516	.36

(c) *Discussion of results.*

(1) The relative order of predictive efficiency for success in both Junior and Senior Certificate is: (1) Arithmetic, (2) Intelligence Test, (3) English Language, (4) English Composition.

This order is the same as that found in the Lancashire experiment. The two experiments are sharply contrasted but yield similar results. The Lancashire experiment was conducted in one school with children known personally to the

writer. The selection examination comprised objective attainment tests with 'new-type' questions. The Northern Ireland experiment used children in many grammar schools and the selection examination used 'traditional' scholastic examinations. Yet, similar results are obtained. The Northern Ireland Arithmetic papers have been subjected to criticism at times, but it is clear from the results obtained that the test has succeeded in its function. It might be argued that this type of Arithmetic paper and Intelligence tests are both good predictors because they are measuring the same thing, but such an argument is refuted by the size of the correlation between Arithmetic and Intelligence. This was only 0.49 and is lower than the corresponding correlation of 0.54 between English Language and Intelligence.

(2) Since both experiments find Arithmetic as the best prediction of success in the grammar school, it may be of interest to examine the nature of the Arithmetic papers in some detail. In the Lancashire experiment the papers set are of the objective type, very similar to the Moray House Arithmetic tests. If Arithmetic papers of a more traditional kind are used, as in Northern Ireland, it is to be expected that the skill of the examiner will play a larger part in the success of the paper. The 1948 Arithmetic paper consisted of 20 questions of a problem type and the time allowed was $1\frac{1}{2}$ hours. Though the questions were of varying standards of difficulty, the candidates were informed that each question carried the same number of marks and also that it might not be necessary to answer all the questions to get full marks. The following selection of questions will give an indication of the type of paper used :

1.—Find the cost of 100 stamps at $2\frac{1}{2}$ d. each.

5.—Look at (a) opposite. $\frac{1}{4}$ is less than $\frac{1}{3}$ so the word 'less' has a line under it. Draw a line under the correct word (greater, equal or less) for (b), (c), (d), (e) and (f).

- (a) $\frac{1}{4}$ is greater than, equal to, less than $\frac{1}{3}$.
- (b) 0.45 is greater than, equal to, less than $\frac{1}{2}$.
- (c) 5.632 is greater than, equal to, less than 17.6×3.2 .
- (d) 1.73 is greater than, equal to, less than $1.245 \div 0.065$.
- (e) $1\frac{1}{2}$ is greater than, equal to, less than $\frac{1}{2} + \frac{1}{3} + \frac{1}{4}$.
- (f) $\frac{1}{3}$ is greater than, equal to, less than half the sum of $\frac{1}{2}$ and $\frac{1}{4}$.

9.—A dealer bought 5 tons of patent fertiliser at £23 7s. 6d. per ton. He sold a quarter of it in bags holding $\frac{1}{4}$ stone each at 15s. 3d. per bag, and the remainder in bags holding $\frac{1}{2}$ stone each at 2s. 5d. per bag. How much did he gain?

13.—A certain machine makes 125 cigarettes per minute. If it is in operation from 8.17 a.m. to 12.28 p.m., and from 1.32 p.m. to 5.58 p.m., how many cigarettes will it make?

17.—A boy arranged his toy soldiers in rows of 12 and had 5 over. If he had arranged them in rows of 9, how many might he have had left over. NOTE.—There are three answers. Give them all.

The examiner's intention in this type of paper is clearly to measure ability as well as attainment. The results of the follow-up experiment would seem to indicate that he has succeeded to a large extent. In assessing the standard of difficulty of this paper, it is well to remember that there tends to be a greater emphasis on Arithmetic in the primary schools of Northern Ireland than in the corresponding English primary schools. It is a universally held belief in Northern Ireland that the standard of Arithmetic at the age of 11+ is higher than in England.

(3) The fact that both follow-up experiments show Arithmetic as the best predictor of success leads one to look for some other possible explanation which might be common to the two experiments. One possibility is that the criteria of success used in both experiments lay more emphasis on science and mathematics than in previous research. In the Northern Ireland criterion, three-eighths of the criterion consists of Mathematics and Science, in the Lancashire experiment since the average mark for all subjects taken in the General Certificate of Education was used, the proportion is somewhat similar. This proportion would seem to the writer to be a reasonable one, but it is higher than some other investigators have used.

(4) The low correlation of English Composition with success in Junior and Senior Certificate is a disappointment to those who advocate the continued use of such an essay. One of the reasons for the low correlations could be the unreliability of the marking of English essays, but no evidence exists, in the experiment, to prove or disprove this hypothesis.

(5) The correlation of 0.74 between the Junior and Senior Certificate is indicative of a substantial change in the order of merit of the candidates in the two examinations. This change is away from the order of merit in the 1948 11+ examination. Thus, the Qualifying Examination is a much better predictor of success in Junior Certificate (taken after three years) than of Senior Certificate (taken after five years).

(6) The most suitable figures to be added to the results in Table I are probably those derived from Junior Certificate since the criterion used was a good one with few limitations. Omitting English Composition these would be :

	Arithmetic	English	Intelligence	No. in sample
Correlation with success	.543	.489	.516	472

V.—CONCLUSIONS.

(1) In Table IV the results obtained from the Lancashire and Northern Ireland follow-up studies are added to those obtained from the survey of previous research. At the foot of Table IV the final average figures (weighted,

TABLE IV
CORRELATIONS WITH SUCCESS.

	Arithmetic	English	Intelligence	Size of Sample
Previous Research284	.355	.422	1582
Lancashire Experiment254	.191	.211	227
Northern Ireland Experiment	.543	.489	.516	472
Final Average340	.369	.423	2281

and computed throughout by using Fisher's Z) are given. Computed, in effect, from a sample of size 2,281, the figures obtained are stable, and would not be greatly altered by any other single experiment. The difference in the size of the English and Intelligence correlation coefficients is significant at the 5 per cent. level, the difference between Arithmetic and English is not quite significant. It is worthy of note, that, of the nine sets of results used to derive the final figures, only one study (Emmett $N=765$) has yielded a difference between English and Intelligence significant at the 5 per cent. level. It is clear that

differences in the relative predictive value are small, but significant if a large enough sample is taken. The final results give support to the view that a verbal intelligence test is the best predictor of success in grammar schools. When it is remembered that Arithmetic and English are always a part of the criterion of success itself, it is quite remarkable that an intelligence test should predict success a little better than attainment tests in Arithmetic and English. The difference in predictive efficiency between English and Arithmetic is small and, in view of the two sets of results contributed by the author, it may be that under post-war conditions the Arithmetic test has increased its predictive efficiency.

(2) Any method of averaging correlations has limitations and the figures given in Table IV should be thought of as a convenient way of summarising a vast amount of research. It is, however, useful to quote that research in another way. Table V shows the order of merit for predictive efficiency of Arithmetic, English and intelligence tests in the nine sets of results. It is a slightly longer, and less statistical way, of expressing the results obtained. From the table it can be seen that five sets of results give an order of merit (1) Intelligence, (2) English, (3) Arithmetic. This evidently is the most usual order; two sets, both due to Stubbins gives the order (1) English, (2) Intelligence, (3) Arithmetic, and finally the two sets of results reported in this article give (1) Arithmetic, (2) Intelligence, (3) English.

TABLE V
ORDERS OF MERIT FOR PREDICTIVE EFFICIENCY.

	Arithmetic	English	Intelligence
Stubbins (1)	3	1	2
Stubbins (2)	3	1	2
McClelland	3	2	1
Emmett	3	2	1
Peel-Rutter	3	2	1
Emmett-Wilmot (1)	3	2	1
Emmett-Wilmot (2)	1	3	2
Lancashire	1	3	2
Northern Ireland			

(3) Whatever the relative merits of the three component parts of the examination, as predictors, it is well established that using Arithmetic, English and Intelligence as a battery of three tests, the correlations obtained are higher, that is to say prediction is more efficient. There would seem to be no valid grounds for eliminating any one of the three usual parts of an 11+ selection examination.

(4) It may be that some results have been omitted from this survey which could have been used. Local authorities may carry out follow-up experiments as routine work without publishing their results. There may also be some useful results contained in unpublished research theses. Any new data would have to be rather startling to alter the obtained averages by very much, but it would be possible, using weighted averages, to add any other result on to the figures obtained in this article.

VI.—SUMMARY.

1.—A survey of published research showed that the relative order of merit for the prediction of success in grammar schools, by the component parts of the

11+ selection examinations, was (1) Intelligence test, (2) English, (3) Arithmetic. Thus, the generally accepted view of educational psychologists, which has been recently questioned, was upheld.

2.—Two follow-up experiments reported by the writer, one in Lancashire and one in Northern Ireland, gave a different order of merit, viz. : (1) Arithmetic, (2) Intelligence test, (3) English.

3.—When the results of these two experiments were added to the results of the survey the order of merit was not changed, but the gap in predictive efficiency of English and Arithmetic was narrowed considerably.

4.—The follow-up experiments in Lancashire and Northern Ireland are the first to publish results showing Arithmetic as the best predictor of grammar school success.

NOTE.

One inevitable defect in follow-up experiments is that by the time they are completed the selection procedures under investigation have changed considerably. In the Lancashire experiment the local authority had introduced a new selection procedure in 1946, and in the Northern Ireland research, 1948 was the first year of operation of the Qualifying Examination. Both authorities have made it clear that they consider their present procedures more efficient than might be indicated by the results presented in this article.

AN INVESTIGATION BEARING ON THE LAWS RELATING TO FEEBLE-MINDEDNESS, AND THEIR APPLICATION

BY F. K. FITZPATRICK

Clinical Psychologist, Botleys Park Hospital, Chertsey, Surrey.

I.—*Introduction.* II.—*Historical background.* III.—*Purpose of investigation.*
IV.—*Selection of patients.* V.—*Investigation.* VI.—*Results.* VII.—*Discussion.*
VIII.—*Summary.* IX.—*References.*

I.—INTRODUCTION.

THIS study was carried out at the suggestion of the Physician Superintendent of Botleys Park Hospital. It was occasioned by the observation of O'Connor and Tizard¹ (1954) that the average I.Q. of young adult feeble-minded patients is more than 70 points. This raises the question whether the working concept of feeble-mindedness has been widened in recent years to include socially inadequate persons who might be better dealt with by other social services, thus freeing valuable hospital beds for undoubted mental defectives of lower grade.

II.—HISTORICAL BACKGROUND.

Feeble-mindedness was recognized as early as 1897, when a Committee was set up to enquire into the existing system for the education of feeble-minded children, "not under the charge of guardians, and not idiots or imbeciles." For the purpose of their enquiry this Committee described the feeble-minded as "children who are capable of earning their living, wholly or partially, in after life," as distinct from idiot and imbecile children who must always be unemployable. This early statement of a social criterion was not incorporated in the definition given in the Elementary Education (Defective and Epileptic) Act, of 1899, which ran on purely educational lines.

The Wood report² (1929), drew attention to this natural tendency for educational legislation to regard mental deficiency as a matter of limited educability, whereas the Mental Deficiency legislation defined it primarily in terms of poor social adaptability. This discrepancy was, however, effectively rectified by the Education Act of 1944. Under it children, fit to attend special schools, were no longer stigmatized by legal certification, when, at a later date, they might be found to be capable of making a successful social adjustment, and so be eligible for 'de-certification.'

The Wood report states that "in the light of all the definitions contained in the Mental Deficiency Acts, and of the best scientific opinion, we have taken the view that, whatever may be the correct legal interpretation of those definitions, the real criterion of mental deficiency is a social one, and that a mentally defective individual, whether child or adult, is one who by reason of incomplete mental development, is incapable of independent social adaptation." This amounts to a re-statement of the original Act of 1913, which is still substantially in force, but with a shift of emphasis to the social aspect. No doubt this would be welcomed by the administrators of the law who saw social incompetency as the immediate reason for certification. Incomplete mental development is at best a vague concept, and without further physiological knowledge is difficult to prove in the case of a sub-cultural defective. Though the intelligence test is a very valuable instrument for assessing the degree of educability, it is con-

siderably less useful in the diagnosis of a social incompetency which may well be due as much to emotional as to cognitive inefficiency.

A new category, the Moral Imbecile (changed in 1927 to Moral Defective) was created by the Act of 1913. This was clearly intended to cover a group of patients who had not given evidence of educational subnormality whilst at school. In other words, their I.Qs. might be expected to approximate to normality. The term 'moral defective' has never been fully accepted, and has largely fallen into disuse, to be replaced by the equally objectionable term 'psychopath' or vaguer 'social problem defective.'

The Act of 1927 changed the wording of two clauses, in order to clarify them: "Existing from birth or from an early age," was now, "existing before the age of eighteen"; and an additional phrase was inserted: "Whether arising from inherent causes or induced by disease or injury." This meant that a young post-encephalitic, for example, exhibiting the characteristic personality defects, could be certified even if his innate intelligence had been normal. Probably these changes did not actually widen the practical scope of the legislation, but they serve to demonstrate that the legal concept of feeble-mindedness is a wide one.

From these brief observations on the history of feeble-mindedness in Britain, it seems clear that it has always been a wide concept, and was intended to be so. And since, in the case of adults, it is recognized to be largely a question of social adaptation, any attempt to define the upper limit solely in terms of an I.Q. would be absurd. Sarason³ (1949) writes: "That the criterion of social adequacy is given far more weight by Tredgold than that of intellectual level is seen by his reference to cases with I.Qs. between 90 and 100 who are considered mentally defective and cases with I.Qs. considerably less than 70 who are not considered defective." Many of those responsible for the certification of these patients would agree with Tredgold, and would welcome an adequate and objective method of assessing social adaptability. Doll's pioneer work in this field, the Vineland Social Maturity Scale is used in this country with children who are already institutionalized, but it cannot claim to be adequate for use with adults who have not been dealt with previously.

III.—PURPOSE OF INVESTIGATION.

The observation, which has been corroborated from more than one hospital, that the average I.Q. of male feeble-minded patients is more than 70 points, could be interpreted in one of two ways: either the patients at the top end of the range are, indeed, defective, within the meaning of the Act, or else the application of the Act has become looser over the course of time. The purpose of this study is to investigate the second of these two possibilities:

(1) By comparing the mental capacity of a group of patients certified twenty or more years ago, with those who are being similarly dealt with at the present time.

(2) By comparing the available data regarding social inadequacy.

It was hoped that some light would be thrown on the suitability of the prevalent methods of diagnosing feeble-mindedness.

IV.—SELECTION OF PATIENTS.

Botleys Park Hospital accommodates mental defectives of both sexes and all ages. On 31st March last, there were 1,669 patients on the books, of whom 964 were male and 705 female. About half of the total are graded as feeble-minded. The catchment area is West Surrey, part of East Hampshire, and the

city of Portsmouth. In the course of twenty-two years there has been a gradual increase in the proportion of males to females, and, principally on the female side, an increase in the proportion of idiots and imbeciles to feeble-minded. Although there has been a change of Physician Superintendent, that has not occasioned any change of policy as regards admissions.

The hospital was first opened to male patients in 1932, so that fixed the extent of the investigation at roughly twenty years. Two groups, each consisting of 40 men between the ages of 16-30, were taken for comparison, one comprising the earliest available admissions and other the most recent.

Of the first 83 men who were admitted at this age, from the beginning of 1932 onward, 25 had been discharged before 1952, 10 had been transferred to other hospitals, 1 had died, and 7 were on licence at too great a distance from the hospital to make testing practicable. The remaining 40 made up the early group. It cannot be claimed that these cases were typical of the intake to this hospital at that time. But it may be argued that the 25 who have been discharged must have been successfully rehabilitated, and in the majority of cases sufficiently intelligent to, at least, contribute to their support. Therefore, it is improbable that they represent the lower end of the I.Q. range.

The second group consisted of the total male intake of this age for the 26 months preceding March 26th, 1954.

V.—INVESTIGATION.

The Wechsler-Bellevue Adult Intelligence Scale (Form 1) was used throughout, with one exception. One of the early group is now inaccessible to testing on account of a supervening psychosis. In his case the Terman and Merrill I.Q. assessed at a child guidance clinic, before admission, was accepted. Two others had deteriorated to imbecile level but not below the lower limit of the scale.

Although the Wechsler Scale is only in process of revision and standardization for use in this country, it has the two advantages of having been designed for use with adults, and of making allowance for deterioration with age. Each year finds its own adult denominator. Wechsler⁴ (1944) claims that "the method enables us to calculate I.Qs. which maintain the same meaning throughout the life of the individual . . . It retains the only important meaning of the I.Q., namely, that of an index of relative brightness."

So, although the magnitude of the scores made by British subjects may prove to be inaccurate, it should still be possible to compare the two groups with a fair degree of confidence even though the early group was tested some twenty years after admission, when they might be expected to have undergone some measurable degree of deterioration. Table I gives the distribution of I.Qs. in the two groups. Table II gives the mean I.Q. and standard deviation of each group, and the result of a test for the significance of the difference of the means.

TABLE I

I.Q.	Old	New
<60	4	9
60-69	13	11
70-79	9	13
80-89	8	7
90-99	5	0
>99	1	0
Total	40	40

TABLE II

	Old	New
Mean I.Q.	74.90	69.00
S.D.	12.35	10.38
t	2.23 (Significant between the 5% and 2% level of probability.)	

The results show a slight but significant superiority on the part of the earlier group. On the other hand, when a comparison was made of such Binet mental ages as were quoted on the certification documents, it was found that the range was of the same order in both groups ($t=1.06$). The question of deterioration need not here be taken into account since the average age of the two groups at that time was 19.2 and 18.6 respectively. Unfortunately, the mental ages quoted on legal documents are not always reliable, and frequently only claim to be approximations. However, if we assume that the two groups were roughly equal in intelligence at the time of certification, the apparent superiority of the earlier group when tested 20 years later needs some explanation. It is possible, but unlikely, that the American adult deteriorates more rapidly than the British, or that defectives deteriorate more slowly than normal people; or it may be that Wechsler's claim that his I.Q. remains constant through life is unjustified. But it seems more probable that defectiveness is generally accompanied by some degree of retardation which diminishes with suitable training. In which case one would expect the recent group to reach the level of the earlier one in twenty years time. A.D.B., and A. M. Clarke⁵ (1954) have observed that there is a tendency for the I.Qs. of young feeble-minded patients to increase, especially when they come from 'social problem' homes.

TABLE III

		Old	New	Total
HOME	Normal	22	20	80
	Broken	9	11	
	None	5	8	
	Not known	4	0	
SCHOOL	Normal	15	18	80
	Normal + Approved	1	3	
	Special	16	15	
	Ineducable	2	4	
	Not known	6	0	
COURT RECORD	Certified under Section 8 or 9	12	20	80
	Certified under Section 6 or 7	11	6	
	Nil	17	14	
OCCUPATION	Wage earning, steady	2	10	80
	Wage earning, many jobs ..	6	10	
	Odd jobs only	15	7	
	In training	1	1	
	Unemployed	9	12	
	Not known	7	0	

The assumption from this study is that as far as this hospital is concerned, the application of the M.D. Acts has not been extended in the past twenty years. Nor does the data regarding social background, set out in Table III, show any striking changes.

The criterion for a normal home was simply that the patient was living with both parents or guardians at the time of certification. 'Broken home' signifies any irregularity in this structure, and 'No home' indicates any form of institutionalization. That a large number received normal education does not mean that their deficiency was not noted at that time. In most cases there is evidence that they were considered backward to some degree, and three were recommended for special schools although they were never admitted. The heading 'special school' includes schools for epileptics and for sense-deprived children. Those described as ineducable were excluded at an early age and attended an occupation centre. Under 'Court Record' all criminal charges have been taken into account whether or not they were immediately antecedent to certification.

The occupational difference (obtained by contrasting all wage-earners with non-wage-earners) is statistically different at the 1 per cent. level (chi-squared). This is easily understood when one recalls the slump of the early thirties when 3,000,000 were out of work. Another interesting point is that the number of the recent group who were certified under Sections 8 or 9 is considerably, though not significantly, greater than that of the older one, although the total of those charged with a criminal offence is only slightly greater. This, no doubt, indicates a change of policy in the course of two decades. (Section 8 and 9 imply that the patient has been convicted of a crime, or is actually in prison when the order is made.)

VII.—DISCUSSION.

This type of enquiry into the past must necessarily be rather inconclusive. The information available is incomplete and perhaps not always reliable. But there is no reason why we should not institute more carefully planned enquiries into the present populations of our hospitals for mental defectives. Surely, it is only by a closer scrutiny of the higher grade patients that we can judge whether or not the law has been too widely interpreted. The actual widening process does not seem to have taken place in recent years, in all probability it occurred in the first quarter of the century, when the science of mental testing was less advanced. Perhaps it is conducive to error. The crucial point is that *there should be no possibility of error in the diagnosis*. If that can be ensured, then it surely follows that the existing M.D. organization is the right one to deal with these patients, and the money which would be needed to enlarge other social services should be spent on providing hostels where the feeble-minded might be trained apart from the defectives of lower grade.

The present method of diagnosis comprises the evidence of an intelligence test, a clinical judgment, and the social and school history. This may be sufficient in the case of children, but as we have already said, the I.Q. is of limited use in the diagnosis of feeble-mindedness in adults. The qualitative information gained from the testing situation may be more useful, but it is often overlooked because it savours of a subjective judgment. The value of the clinical opinion is lessened by the fact that one of the certifying doctors may know little or nothing about deficiency. The social history carries great weight in the diagnosis. But a social history is always relative to the environment. Thus, a confirmed thief, in a criminal family shows no defect of adaptation to the family. There is often a tendency to discredit information tendered by the

patient. Moreover, the social history is not an objective guide; it can, for example, be distorted by parents who wish to dispose of their liability. We have already discussed the need for an adequate social maturity scale, but there is an equally important need for some objective measure of emotional development.

This question of diagnosis is too serious to discuss fully here; however, we might consider one approach, before leaving the subject. Surely, the possibility of error could be minimised by a more co-ordinated effort. If all high-grade and border-line cases could be kept under expert observation in as normal circumstances as possible, for a few weeks before the final step of certification were taken, we might get the information we require. Hostels might be specialized for this purpose, where the subjects could be under constant psychological observation, and in the care of a warden with liberal views on his custodianship. Within this structure it should be possible to build up a system of realistic tests, on the same principle as the procedures used in various other services for selecting personnel.

VIII.—SUMMARY.

1.—Although it is generally assumed that the I.Q. range of feeble-minded patients in England is roughly 50-70 points, this assumption does not correspond to observed fact. Though 50 points may be taken as the lower limit, there is, in practice, no definable upper limit provided evidence can be produced of arrested or incomplete development of mind existing before the age of eighteen.

2.—A brief survey of the relevant legislation seems to show that the legal concept of feeble-mindedness has always been a wide one.

3.—In order to ascertain whether the M.D. Acts have come to be more loosely applied in recent years, in the Counties of Surrey and Hampshire, the following investigation was carried out: Two groups, each of 40 male patients, between the ages of 16 and 30, were compared with regard to I.Q. and such social data as could be obtained. The first group were certified as feeble-minded 20 or more years ago, and the second were the most recent admissions in that grade, to Botleys Park Hospital.

4.—The result of the comparison of I.Qs. showed a small, but significant ($t=2.23$) superiority on the part of the older group. It seemed doubtful whether this evidence of superiority could be accepted. But at all events, there is no evidence that the newer patients are more intelligent than the older ones, so it seems probable that any loosening in the application of the M.D. Acts took place more than twenty years ago.

5.—The social histories showed a significantly (1 per cent. level) greater number of wage-earners amongst the recently certified group. This can be explained by the difference in the industrial conditions of the two eras.

ACKNOWLEDGEMENT.

I am greatly indebted to Dr. J. M. Crawford, Physician Superintendent of Botleys Park Hospital, for suggesting this paper and for his assistance.

IX.—REFERENCES.

- ¹ O'CONNOR, N., and TIZARD, J.: *Brit. Med. J.*, 1, 16 (1954).
- ² WOOD, A. H.: *Report of the Mental Deficiency Committee* (1929).
- ³ SARASON, S. B.: *Psychological Problems in Mental Deficiency* (1949).
- ⁴ WECHSLER, D.: *The Measurement of Adult Intelligence*, p. 35 (1944).
- ⁵ CLARKE, A. D. B., and A. M.: "Cognitive Changes in the Feeble-Minded, *Brit. J. Psych.*, XLV, 3 (1954).

THE INCIDENCE OF NAIL BITING AMONG SCHOOL CHILDREN

By L. B. BIRCH,

Deputy Director, University of Sheffield Institute of Education.

I.—*Introduction.* II.—*Method of the inquiry.* III.—*Results.* IV.—*Conclusions.*
V.—*Summary.*

I.—INTRODUCTION.

At a psychology class, attended by practising teachers, the writer of this account was approached by the head master of a junior school, who said he was very worried to find that almost half of the children in his top class bit their nails, a few severely enough to cause considerable discomfort and even bleeding. Most of the other teachers present said that they, too, had noticed that many of the children in their classes bit their nails, though they had not actually counted the number. They undertook to do so for the next meeting of the class, and the writer promised to find out in the meantime whether there had been any recent surveys carried out on normal school populations which could be used for comparison. He was unable to find anything and when, at the next meeting, all the teachers reported a high proportion of nail biters among the children they taught, he decided to approach Professor Valentine to see whether he knew of any surveys on the topic. Before he had done this, however, he received a letter from Professor Valentine which really began the enquiry which is reported below.

It so happened that Professor Valentine had himself become interested in the incidence of *nail-biting* among children and had searched the literature for results of large-scale surveys which he could use to illustrate a section in his latest book. Having discovered that about 60 per cent. of the boys and rather less than 60 per cent. of the girls between the ages of 8 and 11 in *Chicago*, were reported to be nailbiters, he had, in the absence of any recent English work, begun an enquiry in this country. His first data, from rather small numbers, seemed to indicate that the incidence here was probably not very different from that in the American survey and his letter to the writer was a request for help in the collecting of data from a larger sample in order to verify this indication. When this request was made known to the group of teachers who had raised the question with the writer they at once offered to help, as did members of several other classes when they heard about it. It was clear that results from a reasonably large number of children could be collected fairly easily and when Professor Valentine heard of this he was pleased to leave the enquiry to the writer and, in addition to suggesting a scale for use in recording the degree of nail biting of the children, he also made available the data which had already been collected for him by a number of teachers. These data have not been included with those collected for the writer though they are referred to in other places.

II.—METHOD OF THE INQUIRY.

As has been indicated, the data for this inquiry were mostly collected by teacher volunteers who were attending part-time courses of various kinds in the University of Sheffield Institute of Education. One of the courses at the time was on elementary statistics and, because nearly all the members of this class were men and because this class provided a large number of volunteers, there are rather more boys than girls in the final sample of children.

Altogether, seventy-eight teachers from thirty-nine schools took part in the inquiry; some reported upon only one class, others on more than one and a few on a whole school. They were not selected specially for the task, but, having volunteered, were carefully instructed upon how to scale and record their results. Each teacher was given a printed sheet, giving the following details for scoring.

*Score**Degree of finger-nail biting.*

- 0 No finger bitten—Free margin of nail beyond soft tissue intact on all fingers.
 1+ Mildly bitten—Free margin of one or more nails irregular parts bitten (if any doubt, question child).
 2+ Moderately bitten—Free margin of nail absent, i.e., fingernail bitten to the soft tissue.
 3+ Severely bitten—Nail bitten *beyond* the free edge; nail margin now below the soft tissue border.

Below this schedule were drawings showing exactly what was meant by each of the four degrees of nail-biting. Teachers were asked to assess all the children in any particular class, but were told not to try to obtain the scores of children who were absent. Printed record forms were provided for recording each child's sex, age last birthday, type of school, and nail-biting score. A small group of teachers also agreed to ask all the nail biters in their classes when and under what circumstances the children actually indulged in the biting.

TABLE I
SURVEY OF NAIL BITING AMONG CHILDREN

Age of Children	Numbers of Children at each age			PERCENTAGE OF CHILDREN IN EACH CATEGORY											
				0 (No finger bitten)			1 (Mildly bitten)			2 (Moderately bitten)			3 (Severely bitten)		
	Boys	Girls	Total	Boys	Girls	Boys and Girls tog'th'r	Boys	Girls	Boys and Girls tog'th'r	Boys	Girls	Boys and Girls tog'th'r	Boys	Girls	Boys and Girls tog'th'r
16+	50	42	92	70	68	69	12	14	13	14	9	13	4	9	5
15+	187	158	345	45	71	55	28	10	20	17	12	15	10	7	9
14+	287	243	530	44	53	48	21	22	21	16	11	14	21	13	17
13+	284	232	516	43	53	47	19	15	17	18	18	18	22	15	18
12+	315	231	546	38	52	43	21	15	19	18	18	18	22	16	20
11+	310	262	572	48	49	48	21	19	23	13	16	14	18	16	17
10+	256	122	378	40	53	44	20	16	19	19	14	18	20	17	19
9+	263	107	370	44	62	48	20	14	18	18	9	15	19	15	18
8+	176	61	237	43	52	46	30	8	23	13	18	14	16	21	17
7+	193	83	276	51	61	54	25	18	23	12	10	12	11	12	11
6+	153	106	259	59	51	56	19	18	19	9	16	12	13	15	14
5+	57	45	102	74	51	64	9	18	13	5	25	14	12	7	10
Totals	2531	1692	4223	46	54	49	21	16	19	16	15	15	18	15	17

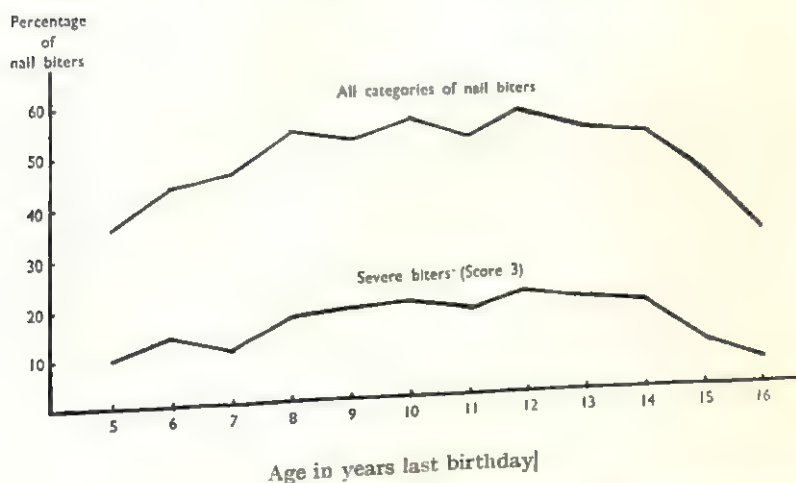
III.—RESULTS.

(a) *In General.*

Table I gives the percentage frequency of nail-biting at various ages of boys, girls and boys and girls taken together. It will be seen that, in the whole sample, roughly half of the children bite their nails to some extent, there being

51 per cent. biters and 49 per cent. non-biters (S.E. = .77). On the whole, boys are more frequently biters than girls, the difference between the 54 per cent. for boys and 46 per cent. for girls being highly significant. This difference is apparent in the figures for all ages between 7 and 15 though, in many cases, the difference is not statistically significant. Below the age of 7, the results seem to show that boys are less frequently nail biters than are girls, unfortunately the numbers here are too small to establish significance. When the results are set out in graphical form, they seem to show a steady increase in the number of nail biters up to the age of 8, followed by a period up to the age of 14 when the numbers remain fairly constant. In fact, the most noticeable feature is the absence of marked peak periods at particular ages, even that at the age of twelve is hardly significantly higher than the percentage found for the whole population.

FIGURE I
PERCENTAGE DISTRIBUTION OF NAILBITERS (BOYS AND GIRLS TOGETHER)



When graphed separately for the sexes, as in Fig. II, it will be seen that there is one marked peak period for nail biting among boys, at the age of 12. The incidence at this age is significantly higher than the percentage found for the whole male population. In view of the connection which some writers hold to exist between anxiety and nail biting, it is interesting to note that there is so little evidence for an increase in the incidence of nail biting at the time of the special place examinations at the age of 10.

When the numbers of children biting their nails are broken down into the three categories of 'slight,' 'moderate' and 'severe' biting, it is found that they are fairly evenly distributed among the three categories.

(b) *Nail biting in different types of school.*

At the secondary age there were, in the sample, children from most types of secondary school, though some, such as those from secondary technical, private and special schools were too few in number to make comparisons worth while. In Table II a comparison is made between the results obtained for grammar school and secondary modern school children between the ages of 11 and 15.

FIGURE II

PERCENTAGE DISTRIBUTION OF NAILBITERS (BOYS AND GIRLS SEPARATELY)

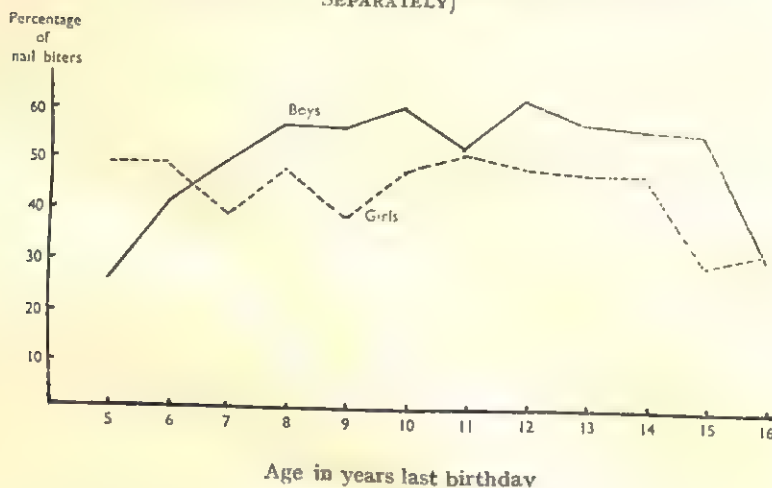


TABLE II

PERCENTAGE OF NAIL BITERS AMONG CHILDREN AGED 11 TO 15 YEARS IN SECONDARY GRAMMAR AND SECONDARY MODERN SCHOOLS

Type of School	Boys	Girls	Boys and Girls Together
Grammar	50 (N = 392)	38 (N = 413)	44 (N = 805)
Secondary Modern ..	56 (N = 413)	49 (N = 595)	53 (N = 1461)

It will be seen that the grammar school children are less frequently nail biters than are secondary modern children and when the boys and girls are considered together, the difference is almost three times the standard error and so can be regarded as statistically significant.

Results were obtained from 156 boys, aged 11 to 15, in three special (E.S.N.) schools, and also from 196 men and 383 women students, aged 18 and 19, from four teacher training colleges. 47 per cent. of the special school children were nail biters, a smaller proportion than was found among grammar school boys of similar age. The difference is not, however, significant.

Of the training college students, 46 per cent. of the men and 35 per cent. of the women students were found to be nail biters, a figure very similar to that found for grammar school children, though here again the standard error with such small numbers is so large that confident conclusions cannot be drawn.

An interesting feature of the results from the training college students was that very few of them scored 3 on the schedule, there being less than 1 per cent. of the women and less than 3 per cent. of the men in this category of 'severe biters.'

(c) *When do children bite their nails?*

473 nail biters of both sexes, aged between 8 and 10, and coming from five junior schools, were asked personally under what circumstances they bit their

nails. Of these more than half (54 per cent.) either said that they did not know when they bit their nails or said they bit them at no special times. 23 per cent. of them said they bit theirs most frequently at the pictures or when watching television; 8 per cent. when reading an exciting book, hearing an exciting story or seeing an exciting incident shown as a film or on television. 11 per cent. said that they tended to bite when they were concentrating at school, especially when doing arithmetic ('hard sums'). 5 per cent. seemed to bite their nails chiefly when they were bored, about half of this group referring to biting in bed. One boy in this group said that he quite often bit his toe nails too. A small group (2 per cent.) referred to worry or 'being grumbled at' as factors connected with their biting.

A similar range of reasons was given to one of Professor Valentine's teachers who collected her data from a large rural school in Worcestershire.

IV.—CONCLUSIONS.

The sample of children used in this enquiry was probably not a random one, even for the industrial area of South Yorkshire, in which all the schools taking part were situated. On the other hand, it seems unlikely that there were any factors systematically influencing the results with reference to the trait being considered. It is probable, therefore, that the figure of 51 per cent. (S.E. .77) of all children being nail biters in some degree is reasonably representative of the district. It should, of course, be remembered that the results do not show how many children give up biting, their numbers being made up by others who start biting late in their school life. If these were taken into consideration, it seems possible that the number of children who bite their nails at some time or other may be considerably above the half which was found in this survey. Further, as a number of teachers pointed out, there are children who, while not being nail biters do bite their hands somewhere else, e.g., the backs of their hands or the skin round their nails. There are also others having the appearance of nail biters whose nails are torn off by hand rather than by biting. In fact, one teacher in a large boys' secondary modern school reported that all but three children in the four classes he taught had some habit of chewing or 'picking' the nails or the skin on the fingers.

No attempt was made in this inquiry to assess the causes of nail biting. Some writers have viewed the condition gravely, regarding it variously as an expression of repressed aggression, emotional maladjustment or anxiety. Hadfield,¹ for example, writes: "The child who cannot let off its aggression against others, lets off its aggression by biting its own nails." English and Pearson write: "... despite the prevalence in the population we must regard it as a sign of poor emotional adjustment."² On the other hand, Barton Hall³ while agreeing that there are cases "in which there is a pronounced element of anxiety and frustration and in which the nails are unusually very severely bitten," considers that in the majority of cases the nail biting "... is an habitual reaction and is otherwise of no particular significance except perhaps for the fact that children with this habit are often of over active, energetic, restless nature."

Since roughly half the children in our schools and a good proportion of the students in some of our training colleges seem to bite their nails, there is obviously a need to find out to what extent there is a connection between nail biting and maladjustment.

¹ HADFIELD, J. A. : *Psychology and Mental Health*, 1950, p. 303.

² ENGLISH, O. S., and PEARSON, G. H. T. : *Emotional Problems of Living*, 1947, p. 196.

BARTON HALL, M. : *Psychiatric Examination of the School Child*, 1947, p. 181.

V.—SUMMARY.

1.—The hands of some 4,000 children in schools in the industrial area of South Yorkshire were examined by their teachers to see which ones bit their nails and to what extent.

2.—The results of the survey show :

- (a) 51 per cent. of all the children seen bite their nails to some extent, 17 per cent. bite them very severely and 15 per cent. moderately severely.
- (b) The incidence of nail biting seems to reach its maximum near the beginning of the junior school period and then to remain steady up to the age of 14, after which there seems to be a falling off. There is a peak period for boys at the age of 12 and for girls at the age of 11, but only the first of these is significantly higher than the mean level of the whole population of the same sex.
- (c) Boys are more frequently nail biters than girls. The difference (54 per cent to 46 per cent.) is significant. Below the age of 7, however, the results show a reversal, with girls the more frequent biters. Numbers here are too small however, for the difference to be significant.
- (d) Grammar school children bite their nails rather less frequently than do secondary modern school children at comparable ages.

SUMMARIES OF RESEARCHES REPORTED IN DEGREE THESES¹

A Test of Group Conformity in the Moral Judgments of Children.

By F. P. BALL

(Abstract of a thesis submitted for the degree of M.Ed. at the University of Manchester, 1954.)

OBJECT.

The investigation sought to compare, in three different ways, the moral judgment of each individual in a class with the collective judgment of the whole class. The three assessments of the degree of conformity of opinion of each individual were then compared with each other, and with independent criteria.

CHILDREN TESTED.

253 boys and 382 girls in their third year at secondary schools and, therefore, aged 13 plus, co-operated in the enquiry.

MATERIAL.

Preliminary investigations were made with 143 boys and 148 girls of similar age. Their judgments of the relative importance of several 'bad deeds' were more clear-cut than similar judgments relating to 'good deeds.' From lists later made of things they considered 'wrong to do,' the ten most frequently named were chosen for the enquiry proper. These were Bullying, Cheating, Cruelty to Animals, Destructiveness, Drinking, Gossip or Tale-telling, Lying, Rudeness, Smoking, Stealing. They were represented, in random order, in three distinct ways. One was as abstract terms, listed above. Another was in pictures, specially drawn by a thirteen-year-old girl. The third way was in sentence-stories, such as "Ron called his grandfather a silly old fool."

METHOD.

One school period gave the investigator sufficient time to present these representations, separately, and with different introductory wording, to one class. The children were asked, without consultation, to number the offences in order of seriousness in each case. They were also asked to name the child in their class regarded by them as 'most average.' Independently, three teachers of each class were asked to assess the behaviour and attitude of each individual as 'Average' or 'Submissive' or 'Individualistic.' These three opinions combined to classify each child on a seven-point scale.

ANALYSIS.

For each class and each method of presentation of the 'offences,' a mean rank was obtained for each offence. Hence, a pooled rank order from 1 to 10 was obtained. The rank difference correlation between this and each individual's stated order was regarded as the score of that individual. Thus, a high score represented a high degree of conformity. For comparison purposes, the scores were grouped as 'under 0.2' and thence upwards in class intervals of 0.1.

Product moment correlations were used to compare the collective scores of each class on the different forms of the enquiry with each other, in pairs, and separately, with the teachers' pooled opinions. Of the 66 coefficients falling under the former description, 9 were significant at the 1 per cent. level, and a further 7 at the 5 per cent. level. There were only 63 correlations with teachers' opinions, one set of teachers having been unable to provide their views, and of these only 2 were significant at the 5 per cent. level.

¹ These Outlines must be submitted through the Head of the Department in which the research was carried out.

Children named as 'average' not less than three times by their class-mates are compared with the pooled teachers' opinion of them in the following table, where teachers' category 1 is the individualist.

Teachers' Category	1	2	3	4	5	6	7
Number of Average Children	6	8	12	18	9	5	3

The scores made on each form of the test by these same 'most average children' are shewn below :

Score20—	.20+	.30+	.40+	.50+	.60+	.70+	.80+	.90+
Stories	7	5	9	8	9	9	14	7	1
Abstractions....	2	3	9	7	10	8	13	12	5
Pictures	7	6	3	8	7	20	13	3	2

DISCUSSION.

It is known that children's attitudes and judgments tend to be specific rather than general, and that they are modified towards a mean when made by children in a group to which they have an awareness of belonging. This is especially the case when comparison is made between judgments of the actions of others and of the child's own actions, or of the child's behaviour. In the pooled rank orders of seriousness, for any one class, there were distinct differences between the results obtained from the different presentations of the data. Yet, for the different classes in any one school, there were frequently revealed the prevailing trends of official instruction. Among individual scores there were a few cases of a high degree of consistency, but many cases of extreme divergence among the results from the different forms.

In the detailed analyses it appears that both individuals and complete classes have greater consistency of opinion shown by the three forms of test when the children are more intelligent than the majority.

It appears probable that this type of test is capable of fairly accurate summary of such public opinion as is voiced, but it is of very limited use as an assessment of the conformity of the individual to the group. In narrower form, by making each item more specific than on this occasion, it could probably be more useful: informal discussions with classes in the concluding minutes of the period used showed that many individuals had, for a variety of reasons, concentrated on some irrelevant aspect of the presentation and based their judgment on that.

BOOK REVIEWS

Adult Education—Why this Apathy?: ERNEST GREEN. (Allen and Unwin, 1953, p. 146, 15s.)

Dr. Ernest Green, for many years redoubtable General Secretary of the Workers' Educational Association, spent the first year or two of his retirement in trying to find out why the masses are not in the classes ("adult education" and the W.E.A. are more or less synonymous for him). This was a desperate enquiry from the start. It may be possible to enumerate the motives which take people into adult education, but who could ever finish enumerating the motives of 20,000,000 absentees? Ask the man in the street why he doesn't join the W.E.A. and he will say: "Why should I? What is it?" thus presenting a chance to say something sensible and perhaps useful to him. All the best parts of this book can be regarded as answers to *this* question, but the rest is shadowy and unconvincing. Dr. Green went to work with questionnaires; he had replies from 1,011 people active in adult education, and 376 from people apparently not interested. But it is clear that those 376 were not 'apathetic' persons at all; indeed, 140 of them had actually had some contact with adult education. What should have been the real subjects of the enquiry were people not likely to answer questionnaires. The basis of the investigation does not at any point seem to have been clearly formulated, and no really representative sample was found, let alone investigated. However, 448 W.E.A. students worked in 49 study groups at a syllabus which Dr. Green devised for them and produced between them a series of unpretentious and sensible observations, opinions, and judgments, well worth collecting and worthily representing the admirable seriousness of the W.E.A. They offer no statistics, but they account for Dr. Green's deep faith in the Association he served so well. From them you may readily see what the faithful find in the W.E.A.; from Dr. Green himself you may learn something of what the W.E.A. could do to improve its work; but neither from them or him will you get any satisfactory answer, should you feel the need of one, to the question, "Why this apathy?"

R. D. WALLER.

Aggression and its Interpretation: LYDIA JACKSON. (Methuen, 1954, 21s., pp. 5 + 237.)

This book is divided into four parts. In Part I the author discusses various concepts of aggression and the part played by aggression in the causation of delinquency; in Part II she describes three projection studies carried out with various groups of children, normal, neurotic and delinquent; Part III is devoted to clinical studies of aggression in children referred to her for psychological treatment, while in Part IV she attempts an interpretation of results within the framework of a theory of aggression.

The author accepts aggression as a natural mode of human behaviour. "In so far as it is a form of self-assertion and is thus one of the ways to self realisation, aggressive behaviour is a *sine qua non* of individual survival" (p. 27). The problem is then as to why some children direct this aggression into socially acceptable modes of behaviour while others, among whom she would place both the neurotic and the delinquent, do not. She contrasts 'positive' and 'negative' forms of aggression and states her view that "both the neuroses and delinquency are forms of 'negative' aggression, and that the principal external difference between them is that in delinquency aggression is transferred from the family to society at large, whereas in neurosis it is confined to the family group. Many neurotic symptoms have a strong aggressive component, and neurotic behaviour disrupts family life just as delinquent behaviour disrupts the life of a larger group" (p. 49). Whether one accepts the self-assertive impulse and aggressive behaviour as co-terminous or not, one may query the distinction here drawn between the neurotic and delinquent child, for the latter may present neurotic features also. The author's own examples of juvenile thieves whose delinquency appears as much directed against family figures as against society may also lead one to doubt her general contention. And it is certainly true that many neurotic children present no symptoms with a strong aggressive component. In her general conclusions she returns to this distinction, stressing here

the principal internal and etiological differences between the two modes of behaviour. In the delinquent child, infantile strivings have been inadequately handled by a rejecting mother and have become projected on to the outside world "from which he tried to wrest the satisfactions his family denied him; he is prematurely and superficially detached from the family" (p. 218). The neurotic child's strivings are centripetal, not centrifugal. He cannot free himself from his family, and this is mainly, if not wholly, determined by early maternal attitudes of ambivalence with only partial or intermittent rejection.

The evidence on which her conclusions are based falls into two types—experimental studies and clinical observations.

Her experimental studies use projection as a means of eliciting fantasies which throw light on the personality and attitudes of her subjects. Her first investigation was directed towards the presence of sado-masochistic trends in delinquent adolescent girls. Using a set of six pictures in black and white, she found differences in sado-masochistic fantasies between normal girls drawn from various girls' clubs and delinquent girls, aged 12-19 years, some inmates of a hostel for unbilletable girls and others inmates of a Home for Moral Welfare. There were 20 girls in each group. She found greater frequency of sado-masochistic stories and stories with unhappy endings in the delinquent group as compared with the control group. "Suffering and cruelty depicted in themselves had a compensatory function and provided satisfaction for some perhaps deeply unconscious need. Their delinquency and their sado-masochism originated in some earlier experiences of which their fantasies were a reflection" (p. 58).

In order to arrive at a clearer picture of what these needs might be and how they might have originated, the author devised a test of family attitudes, using a set of pictures designed to elicit fantasies about various types of family relationships, and appropriate for use with younger children aged 6-12 years. This test was given to three groups of children, 40 normal, 40 neurotic and 30 delinquent, matched for age, sex, intelligence quotient, but apparently not for social class.

Details of the matching procedure are not given. We are told that there were only ten girls in the delinquent group, so it is difficult to see how matching for sex was carried out in the three groups. It is on the results of this investigation that Miss Jackson's later discussion is mainly based.

She divides the replies to the pictures given into various types of categories, e.g., for Picture O (woman with cot in which is, presumably, a baby), nine types of responses were distinguished, for Picture I, sixteen types and so on. She compared the frequency of each type for each picture as given by the three groups, using "Fisher's Chi-squared method" (which, in fact, originated in this form with Pearson), and gives an example of one such calculation on page 86. Unfortunately, the procedure applied to the data in this example (response to Picture O—Woman is baby's mother) appears to be wholly incorrect. If, as stated there, the number of 'Yes' responses by normal subjects is taken to represent the expected frequency, the appropriate test ratio is that of the difference in proportion of the 'Yes' responses by normal and neurotic subjects to the relevant standard error. This method, when applied to the example given, indicates a highly significant difference, not, as is stated, a difference that "cannot be regarded as significant."

Mrs. Jackson's formula differs from this and also from the accepted procedure of applying the chi-squared test to the data given, which in the example, would also indicate a significant difference. The formula the author uses does not take account of the size of the total sample, hence her procedure would give the same results no matter what number of children were tested.

Of the results presented here, therefore, one is certainly invalid and Mrs. Jackson's use of an incorrect method vitiates her conclusions. Although such differences or lack of differences between the three groups may, in fact, exist, their existence has not been established. Short of re-examining the original data, we have no means of knowing whether her conclusions are valid or not. This is a pity, since her studies present much of interest, not only to those concerned with the personality of the difficult child, but to those who use projective techniques in the study of personality.

GERTRUDE KEIR.

Annual Review of Psychology, Vol. 5, 1954: Editor C. P. STONE; Ass. Editor: Q. MCNEMAR. (Annual Reviews Inc., Stanford, California, pp. 448, \$7.00.)

This series continues to be of great value for the advanced student and for researchers on individual topics, though most articles are concerned with other aspects of psychology than child or educational psychology. The authors of the various sections include the British contributors, Professor R. W. Russell (on Comparative Psychology), J. Maxwell (on Educational Psychology), and A. Heron (on Industrial Psychology). Of special interest to readers of this journal is the article on Child Psychology by H. L. Koch (though it tends rather to be a series of brief outlines of papers) and most of all the admirably critical article by J. Maxwell.

C.W.V.

The Human Factor in the Army—Some Applications of Psychology to Training, Selection, Morale and Discipline: C. W. VALENTINE. (Aldershot: Gale and Polden, 2nd edit., 1954, pp. 106, 3s. 6d.)

First published in 1943, at a time of considerable official concern for efficiency and speed in methods of training, the new and revised edition of this little book is none the less welcome and apposite to-day when the nation is maintaining the largest peace-time army of her history. "All training is training for war" is an old military maxim while one of newer currency is "All training is character training."

The first edition is well known to Army officers and its second edition will certainly find a still wider circle of readers among N.C.Os. who take "Principles and Methods of Instruction" for their examination for the Army Certificate of Education. For the latter, the chapters on "Mental Training," "Discipline," and "Mental Health" are particularly valuable, distinguished as they are by Professor Valentine's uncommon commonsense, his sympathetic as well as enlightened observation of the military scene, and his engaging simplicity of style and treatment—with a complete absence of the verbal ectoplasm of more pretentious works.

The prevalence of major errors in training, he rightly remarks, has decreased, but will there not always be the human factor to reckon with in the most enlightened of armies? Much that he says is by now accepted doctrine embodied in official manuals, but this in itself is evidence of the influence exerted by the author and his fellow psychologists who advised the Director of Military Training during the war.

The new edition will not, it is hoped, escape the attention of those responsible for training schemes outside the Services, in particular in the nationalized corporations with their increasing complexities of hierarchy and function. The principles enunciated are just as valid in a civilian context.

H. L. WILLIAMS.

Early Leaving. A report of the Central Advisory Council for Education. (England.) (London: Her Majesty's Stationery Office, 1954, pp. 99, 3s. 6d.)

This report deals with the question of leaving at different ages from maintained and direct-grant grammar schools. It is important both for its clearly presented statistical data and sensible suggestions, and for the psychological questions which it raises. The data was obtained from questionnaires completed by the heads of a representative sample of schools and also by a non-representative sample of former grammar school pupils. In addition, the Council had informal consultations with local education authority representatives, head teachers, etc., and these consultations presumably account for the suggestions on helpful methods of school organisations, for the report states that the evidence from the questionnaire throws no light on this matter. This suggests that the questionnaire may have been at fault: that the right questions were not asked. However, what does stand out very clearly, if not with startling newness, is the difficulty of selection at 11 plus, and the importance—increasing with age—of the influence of the social background on the intellectual achievement of the pupils. The Council strongly recommends an enquiry into the influence of home background. (One has already been reported in this *Journal*). The Council does, however, also state that "Success can be achieved by a good school even in very difficult conditions." More factual evidence about these

'good' schools would be welcome. It should be useful to know why it is apparently advantageous to be a 'late-transfer' rather than a borderline entrant, how pupils' attitudes to school are influenced by competitively defined positions in school groups, how far the stable tradition of a school and its academic successes can compensate for the absence of special courses for the slower pupils.

Many further lines of investigation and experimentation suggest themselves on reading this report and are, indeed, required to tackle the problem of the intake from social group V of "six failures in order to secure one success." M. COLLINS.

OTHER PUBLICATIONS RECEIVED

The mention of a book in this list does not preclude a later review.

- Education in the New Age*: ALICE A. BAILEY. (Lucis Press, pp. 174, 15s.)
- Dynamics of Groups at Work*: HERBERT A. THELEN. (Univ. of Chicago Press, pp. 370, 45s.)
- The Study of Personality: Book of Readings; Commentary and Compilation*: HOWARD BRAND. (Chapman and Hall, pp. 581, 48s.)
- An Introduction to Mental Measurement and its Applications*: C. A. RICHARDSON. (Longmans, Green, pp. 102, 8s. 6d.)
- Child Psychology—Fourth Edition*: ARTHUR T. JERSILD. (Staples, pp. 676, 42s.)
- Nature's Second Sun*: DONALD MCLEAN. (Heinemann, pp. 130, 10s. 6d.)
- Numbers in Colour*: G. CUISENAIRE and C. GATTEGNO. (Heinemann, pp. 39, 4s. 6d.)
- Dog Stories, in Basic Vocabulary*: E. W. DOLCH and M. P. DOLCH. (Garrard Press, Champaign, Illinois, pp. 169, \$2.)
- A Dictionary of Pastoral Psychology*: VERGILIUS FERM. (Philosophical Library, New York, pp. 336, \$6.)
- A Modern Introduction to Psychology—Fourth Edition*: REX and MARGARET KNIGHT. (Univ. Tutorial Press, Ltd., pp. 242, 8s. 6d.)
- Treasure Ahead*: IAN SERRAILLIER. (Heinemann, pp. 138, 4s. 6d.)
- To Please You*: RONALD RIDOUT. (Heinemann, pp. 153, 4s. 6d.)
- Workbook*: RONALD RIDOUT. (Heinemann, pp. 32, 1s. 2d.)
- Fourth Steps to Drama*: G. H. HOLROYD. (Macmillan, pp. 138, 3s. 9d.)
- Learning Theory, Personality Theory and Clinical Research*: THE KENTUCKY SYMPOSIUM. (Chapman and Hall, pp. 164, 28s.)
- The Art of Teaching English as a Living Language*: I. MORRIS. (Macmillan, pp. 170, 5s.)
- Individualisation en Education Physique*: PAUL FRANKARD and DENISE WALTKEIERS. (E. Nauwelaerts, Louvain, pp. 115, frs. 47.)
- A Study in the Epidemiology of Health*: THE KENT PAEDIATRIC SOCIETY. (Health Department, Bexleyheath, Kent, pp. 88. No price given.)
- Growing up in the City*: J. B. MAYS. (Univ. Press, Liverpool, pp. 216, 17s. 6d.)
- The Handicapped Child—Letters to Parents, I*: DR. KARL KONIG. (New Knowledge Books, pp. 19, 1s. 6d.)
- Factorial Analysis for Non-Mathematicians*: C. J. ADCOCK. (Melbourne Univ. Press, pp. 88, 17s. 6d.)
- Die Tests in der Klinischen Psychologie*: ERICH STERN. (Rascher, pp. 418, frs. 28.10.)
- Researches and Studies*: No. 11, Jan., 1955. (Univ. of Leeds Institute of Education, pp. 99, 3s. 6d.)
- Notes on Learning History*: T. C. EDWARDS. (Blackwell, pp. 16, 1s.)
- Mathematics Notes*. (Blackwell, pp. 16, 1s.)
- Notes on French*. (Blackwell, pp. 24, 1s.)
- Notes on Latin*: R. F. GLOVER. (Blackwell, pp. 20., 1s.)
- Notes on English*. (Blackwell, pp. 16, 1s.)
- Physics Notes*. (Blackwell, pp. 20., 1s.)
- Text Materials in Modern Education*: Editor L. J. CRONBACH. (Univ. of Illinois Press, pp. 216, \$2.50.)

PSYCHOLOGY AND THE TEACHING OF SCIENCE

By E. A. PEEL

Professor of Education, Birmingham University.

- I.—*Science education.* II.—*Scientific thinking.* III.—*Scientific aptitude.*
IV.—*Teaching and learning science.* V.—*Summary.*

I.—SCIENCE EDUCATION.

SCIENCE was one of the last topics to win a place in the school curriculum and the pioneer protagonists for its inclusion in school studies saw it as the core of a new culture. Since the days of Spencer, T. H. Huxley, Tyndall, and Galton, the threefold justification for teaching science on grounds of usefulness, cultural value and training the mind has been regularly restated.¹ Of these three primary grounds for the inclusion of science in the school curriculum, the claim that it trains the intellect and powers of observation is of greatest interest to the psychologist. The disciplinary value of the subject loomed largest in the writings of the last half of the nineteenth century, and it is clear that whatever the pioneers had wished for, science soon had to justify its inclusion on the same grounds as Latin or Mathematics. Research by Thorndike, Sleight and Winch caused the theorists to abandon the disciplinary claims, at least in their writings, but practice in the grammar schools seems always to belie this change, for implied in this practice is the assumption that science, together with the other grammar school subjects, trains for something.

Nowadays, we are, perhaps, more concerned with problems of motivation, interest and activity which, as we now know, provide the conditions for learning in one field to be transferred to another. Here again the fundamental suggestions were made long ago, this time by Armstrong when he deplored the didactic methods then used to teach science and recommended the *heuristic* method which makes science an active study fully motivating the learner. There are now many methods possible for teaching science, some more suited to particular sciences than others.² The general tendency is for both reformers and humbler practitioners to pay more attention to the individual. There is widespread feeling that the interests and powers of the individual pupil tend to be neglected when the class is taught as the unit. Projects also allow for a freer yet cohesive and constructive relationship between individuals and the group of which they are members. These tendencies are both consonant with the educational psychologist's interest in the educative processes in the individual child. However, there are still signs of differences between precept and practice, for a recent investigator³ reported that 40 per cent. of a large representative sample of grammar school science teachers had definitions and principles precisely worded and "learnt by heart." He also noted that "instructions for practical work are rarely given as problems for simple research, but usually as carefully worded directions concerned with nearly every stage of the operation."

¹ See, for example, *The Teaching of Science in Secondary Schools*, I.A.A.M. and S.M.A., pp. 10-11. (John Murray, 1947.)

² *The Teaching of Science in Secondary Schools*, Ch. V. (John Murray, 1947.)

³ RAMSAY, M. P.: *A Psychological Analysis of the Aims of Science Teaching in Secondary Grammar Schools*. (M.Sc., London, 1950.)

Similar changes to those revealed in this country have been noted in science teaching in America where a shift is reported¹ from teaching scientific facts as ends in themselves and an emphasis on formal discipline to the use of facts "as a means of teaching functional understanding of principles of science," to the inculcation of scientific attitudes and to the bearing of scientific method on problem solving.

My purpose is to discuss the bearings of psychology on the teaching of science and in order to do this, I shall ask myself two questions:

To what extent are the findings of psychology relevant to school science work?

How far has progress in the teaching of science resulted at least in part from the psychologist's results and theories?

These questions will be considered in relation to the specific fields of scientific thinking, scientific aptitudes and science teaching and learning, but before we pass to these topics, certain general observations can be made with regard to the second question.

The influence of psychology seems in the past to have acted mainly in a very broad way through catch phrases or very generalised statements such as *formal discipline* (hardly a purely psychological influence, for it was part of educational philosophy long before there was an educational psychology) or the complete *act* of thought as described by Dewey.² It is in this general sense that the methods of science teachers may have reflected contemporary psychological views and not in the more precise way that should be possible by bringing, say, the findings of the gestalt psychologists Wertheimer and Duncker, on problem solving and creative thinking to the notice of the class teacher.

There is surprisingly little research on science teaching that is explicitly designed to test educational practice against psychological theory and experimental result. Thus, in one review of thirty recent researches there is scarcely any mention of psychological topics.³ The same is true of the list of some twenty researches selected by Curtis⁴ as "milestones of research" in the teaching of science in the U.S.A. since 1904. Eight are devoted to learning, but only half of these appear to refer to psychological standards.

When we turn to British work we find that most of the higher degree research on science education that links with psychology is devoted to the problem of scientific abilities, aptitudes, attitudes and interests. Only a few researches refer to learning problems in the psychological setting. For example, in Mrs. Blackwell's three lists⁵ there are twenty-three researches on scientific abilities, whereas there are eighty-six on teaching and learning science of which less than half-a-dozen appear to deal with the bearings of psychology on classroom work. In the sections on the psychology of learning and transfer of training there are reported thirty researches of which perhaps a half have an indirect bearing on the problem of science teaching. A glance through the back numbers of this *Journal* also confirms how rarely the findings of psychology are linked with science teaching.

These observations seem to suggest that the most fruitful method of discussion would be to go straight on to the first question as to what extent psychology is relevant to science teaching.

¹ Faculty of Sch. of Education: *The Improvement of Classroom Instruction*, pp. 11-12. Ann Arbor, 1946. (Univ. of Michigan.)

² The influence of Herbartian psychology was more detailed.

³ Education Dept. Univ. of Chicago, : *The School Review*, Feb., 1954.

⁴ CURTIS, F. D. : "Milestones of Research in the Teaching of Science," *J. of Educ. Res.*, Vol. XLIV No. 3, pp. 161-178, Nov., 1950.

⁵ BLACKWELL, A. M. : *Researches in Education and Educational Psychology*, 1st, 2nd lists and Supplement 1. (National Foundation for Educational Research.)

II.—SCIENTIFIC THINKING.

Scientific thinking is largely a matter of good thinking habits.¹ Bacon drew attention to this centuries ago in his list of 'idola scholarum.' He was so concerned with the dangers of assuming the wrong hypotheses that he dispensed too readily with the necessary assumptions with which the scientist has to start. Whilst insisting that induction is the key to scientific progress, he was, however, critical of induction by mere enumeration. There are two extremes in the opinions on the nature of scientific thought and action. First, there is the positivist view that lays nearly all the emphasis upon the collection of data, followed by classification and induction by "the creative imagination" of laws equally valid for "all normally constituted beings." Comte² founded the movement that insists that experience and action form the basis of scientific law and it was taken up by Mach,³ Spencer,⁴ Darwin,⁵ and Armstrong⁶ in their respective rôles as physicist, populariser, biologist and teacher. The work of Galton in the anthropological and psychological is also characterised by the positive spirit. At the other extreme is the view that denies the naive simplification of positivism. The progress of modern science shows how essential it is to have initial hypotheses and much of the contemporary writing on the nature of science is concerned with this point. Poincaré,⁷ Burt⁸, Russell,⁹ Eddington,¹⁰ and Whittaker¹¹ all stress the importance of hypotheses in the progress of modern science.

These extremes represent to some extent the processes that go on in the pupil and the psychology of creative thinking would seem to have something to contribute to the topic. The common ground of psychology and scientific thinking is seen in Piaget's latest work in an English translation¹² where he traces the development of the child's concept of matter, space, time and causality and on the side of the scientist the writing of Conant¹³ is full of suggestive ideas for the psychologist. First there must be hypotheses, then there is the progress of the scientist "through thickets of erroneous observations, misleading generalisations, inadequate formulations and uncommon prejudice" (¹¹, p. 15) and later we read that "the test of a new idea is, therefore, not only its success in correlating the then known facts, but much more its success or failure in stimulating further experimentation or observation" (¹¹, p. 24). A theory is overthrown by a better theory and never merely by contradictory facts.

Two psychologists have studied similar problems in students and children, Duncker¹⁴ on the mechanism of productive thinking and problem solving and Piaget¹⁵ on the development of children's reasoning and associated powers.

¹ PEEL, E. A.: "Qu'est-ce que l'Aptitude Scientifique?" *La Revue Internationale de Psycho-Pedagogie*, Vol. III, July, 1955, p. 364.

² COMTE, A.: *A Discourse on the Positive Spirit*.

³ MACH, E.: *The Science of Mechanics*.

⁴ SPENCER, H.: *Education*.

⁵ DARWIN, C.: *Origin of the Species*.

⁶ ARMSTRONG, H. E.: *Teaching of Scientific Method*.

⁷ POINCARÉ, H.: *Science and Hypothesis*.

⁸ BURTT, E. A.: *Metaphysical Foundations of Modern Science*.

⁹ RUSSELL, B.: *Analysis of Matter*.

¹⁰ EDDINGTON, A. S.: *The Nature of the Physical World*.

¹¹ WHITTAKER, E. T.: *From Euclid to Eddington*.

¹² PIAGET, J.: *The Child's Construction of Reality*.

¹³ CONANT, J. B.: *On Understanding Science*, 1947.

¹⁴ DUNCKER, K.: "On Problem Solving," *Psych. Mono.*, 58, 5, 1945. See also HUM-

PHREY, G.: *Thinking*.

¹⁵ PIAGET, J.: *Logic and Psychology*.

PIAGET, J.: *The Child's Conception of Physical Causality*.

Duncker traced the course of a person's attempt to solve a problem. A problem arises when the person has a goal but does not know how this goal is to be reached. So there was the problem of the inoperable tumour and rays which destroy organic tissue at sufficient intensity. How can the tumour be destroyed without injury to the healthy tissue? There was also the problem, for the non-physicists, of the compensating pendulum. Lastly, he used several intriguing mathematical problems such as why is it that all six figured numbers of the type *abc, abc*, such as 159, 159 or 842, 842 are divisible by 13? The subjects of the experiments were asked simply to "think out loud" as they attempted to solve the problem, reporting all attempts no matter how irrelevant they were felt to be. The recorded statements gave evidence that progress to the final solution is by a sequence of restatements of the problem. These reformulations tend to become more specific and less general. For example, one such restatement was as follows: How to make the rays more concentrated in the tumour only? This statement is more specifically directed towards final solution than the original formulation of the problem. These statements in prospect have the appearance of problems to the person making them and in retrospect the quality of solutions.

We may note at this stage how similar these situations are to those occurring in every science lesson. Duncker's work suggests that we should allow the correct solution or interpretation of experimental data to emerge organically from earlier attempts. In fact, he demonstrates the superiority of the organic solution over the mechanical method of *rule* and *application* and his work, therefore, lends support to all those who have opposed the didactic method of teaching science.

There has not been much recent research along these lines in Great Britain but one small investigation merits comment. Kyle¹ tried to find whether and how pupils discovered Archimedes principle experimentally. He showed that abler pupils tend to progress by *hindsight*. They tend to jump to the end of certain phases and then to return to fill in the blanks. Thus, when they had to find the volume of an object (by displacement), the abler pupils usually immersed the object first and *then* found the volume of liquid before and after immersion. A problem only becomes *real* for a person when he has some rudimentary foresight of a tentative solution. Thus, we are told² that Einstein usually had a "definite vision of possible solution" of any problem he was engaged upon. Kyle also showed that a "doing" group went further towards a solution than a "thinking" group. All pupils failed to link "upthrust" with volume and weight of matter displaced.

Duncker also investigated the effects of the phrasing of a problem and the materials used in practical problems in the *direction* of the solution. He found that materials usually used in other situations tended to become fixed and, therefore, halted progress. Thus, in one experiment students had to fasten small platforms to a door in order to hold candles. There were several objects available which included three match boxes which in one case were filled with material, thereby being *fixed* or pre-utilised as boxes. In the other case they were presented open and empty (not fixed). The group of subjects working in the latter circumstances were 100 per cent. successful in solving the practical problem whilst the former group were only 42 per cent. successful. Problem solving depends on how quickly elements are recognised as solution material as opposed to conflict elements.

¹ KYLE, T.: *An Investigation of the Thought Processes of a Group of Fourteen-year-olds during the Solving of a Scientific Problem.* (Unpub. Ed. B. Thesis, Glasgow, 1950.)

² HARDING, R. E. M.: *An Anatomy of Inspiration*, 1940.

There have been only few researches on practical physics but those referred to by Ward¹ and his own results confirm the relevance of Duncker's work and the small study by Kyle for the problem of practical work in physics.

Finally, we should ask where Piaget's extensive inquiries bear on the problem of scientific thinking. Mainly in connection with the way in which the scientific concepts emerge in the thinking of the pupils and in the development of the operation of thought. Piaget emphasises the operational nature of thinking. Thinking is an act developed through the phases of sensori-motor responses, pre-operational thought, concrete operations and finally formal operations.² The difference between the last two phases is shown in a physics experiment such as determining the relationship between the length, swing, and weight of a pendulum. Children at the concrete stage vary the factors in a haphazard way, classify and order, and thence set up correspondences. Pupils of twelve to fifteen, however, formulate after a few trials "all the possible hypotheses concerning the operative factors and then arrange their experiments as a function of these factors." (¹, p. 19). Operations emerge fully internalised in thought, reversible, and combined in structured wholes in the formal propositions of mature intelligent beings.

Both by his stressing of the operational nature of thinking and by his inclusion of inductive and deductive elements in his scheme of development, Piaget has a system which is very suggestive for discussions of scientific thinking.

III.—SCIENTIFIC APTITUDE.

The question most commonly asked by the layman is whether there is a special aptitude or ability for science. Is something involved in learning science in addition to general mental ability? If so, where does it stop? Does it include mathematics and other parts of the grammar school curriculum?

On his side the psychologist is interested first in demonstrating the existence of special abilities and then in attempting to identify them in terms of his tests. The method of correlation and factor analysis is used for this purpose. Sometimes research workers devise tests entirely within the field of a school subject and then attempt to show what parts of the subject group together as, for example, in research by Pawley³ on the abilities involved in learning school chemistry. Another relevant kind of inquiry is to conduct a factorial analysis of all grammar school subjects in order to find the special groupings of subjects. Lastly there are researches concerned with devising tests which *predict* scientific aptitude. These tests are often based on factorial studies, but are used in follow-up studies at the end of which tests, assessments or examinations are given in the school subject to be predicted. Such researches are concerned with aptitudes in the narrow technical sense.

Let us first mention research on the relation of science to other school subjects and then discuss the results of researches on science ability and lastly, deal with tests of scientific aptitude.

A fairly recent research on the first topic was carried out by Ellis⁴ who wished to find a psychological basis for the organisation of grammar school courses. His thesis also contains a comprehensive account of earlier work in this field. He analysed the school certificate marks of 200 boys into their factors. He found

¹ WARD, A. E. H.: *A Study of Practical Tests in Physics at VIth Form Level* (Unpub. London M.A. Thesis, 1954.)

² PIAGET, J.: *Logic and Psychology*, Chapter II.

³ PAWLEY, G. C.: *A Factorial Analysis of the Abilities involved in the learning of School Chemistry*, M.A., London, 1937.

⁴ ELLIS, A. G.: *The Relations between Grammar School Subjects*, M.Ed., Durham, 1948.

three factors of which the "first factor had a marked effect on history, geography and English literature and a very small effect on the rest (zero on mathematics); the second was clearly connected with physics, chemistry, mathematics and geography with small effect on the rest (zero on English Language); the third had considerable influence on English Language and French, while Mathematics, Physics, Chemistry and English Literature were all moderately affected (zero on Art)."

Such investigations of the groupings of similar grammar school subjects do not help a great deal with the analysis of science ability, but at least they confirm the common elements in attainment in school science and the elements common both to science and other subjects. A more penetrating analysis is possible in the second type of inquiry, that is, into the nature of science ability by the use of psychological and educational tests.

Several of the researches on this topic were carried out at London and I shall take three of these as examples. First, there is the factorial analysis carried out by Berridge¹ on ability in school physics. He gave a battery of verbal, reasoning, memory, spatial and practical tests and included also school marks in heat, mechanics and hydrostatics. He identified factors of abstract reasoning, memory, use of words, number, induction and spatial reasoning. Induction and reasoning, spatial, abstract and verbal correlated most highly with school physics. These factors were arrived at by rotating the original centroid factors. The unrotated 'g' factor gave highest loadings in hydrostatics and mechanics and lowest loadings in the Vincent Mechanical Models test. Khan² devised tests of accuracy of observation, definition, classification and interpretation and also tests of application, generalisation, planning of experiments and resourcefulness, all in science. As a result of factorial analysis he identified three factors entering into school science: general intellectual power, verbal reasoning and visual imagery. He took the tests of accuracy of observation and definition and the test of application of principles in new situations in order to predict achievement in technical and grammar school science and obtained a workable, but not large maximum correlation (.47). Jog³ used a battery of verbal, non-verbal, spatial and practical tests of ability, together with attainment tests in arithmetic, algebra and physics and tests of persistence. He identified the following factors which enter into attainment in physics: general intelligence 'g,' visuo-mechanical ability and industry (persistence). The extent to which these factors predict attainment in physics is, however, not large. Two other pieces of work not primarily factorial may be noted before we sum up the results of factorial research. The first is that carried out by Young⁴ who devised tests of inferences involving school science subjects and then related them to tests of intelligence and school attainments. Each test consisted of the description of an experimental set-up together with five inferences as to the results of the experiment. The pupils had to mark each of these statements as true, false or 'could not say owing to insufficient evidence.' Young found that intelligence was the major factor in the ability to draw inferences from science experiments but that his test of inference correlated almost equally with attainment in Science,

¹ BERRIDGE, A. R.: *A Factorial Analysis of Ability in School Physics*, M.A., London, 1948.

² KHAN, M. A.: *A Study of Ability and Interest of Secondary School Pupils in Science*, M.A., London, 1951.

³ JOG, R. N.: *A Study of the Factors underlying the Ability of Grammar School Boys to Learn Physics at the G.C.E. Stage*, M.A., London, 1955.

⁴ YOUNG, J. E.: *The Ability to Draw Inferences from Science Experiments*, Ed.B., Glasgow, 1949.

English and Mathematics. Furthermore, the correlation was not high ($\cdot 30$). The other study is the work by Ward¹ which has been referred to. He was primarily concerned with the efficiency of different kinds of practical physics examinations, but found correlations between practical and theoretical physics at the sixth form level which far exceed those found in the researches quoted. Evidently, a practical factor as measured by these examinations enters very intimately into physics attainment. What is practical ability in physics? There is first the "grasp of the problem; then there is the ability to plan the series of movements necessary to bring about the change; finally, there is the ability to execute the movements necessary" (², p. 187).

What is the outcome of these factorial studies? It is clear that they do not lend much support to the view that there is a well marked group factor or well-defined special ability in science. A clearly defined association emerges between scientific aptitude and the general-verbal factor of intelligence and reasoning factors, but the evidence for practical and spatial factors is rather inconsistent and the case for factors of imagination and memory needs also to be verified further. The overall conclusions reached from this brief survey support Vernon's interpretation of mental factors for science.² He suggests that factors of scientific ability cut across the broad verbal-educational and kinaesthetic-motor classification (p. 28) and in his diagrammatic scheme 'science' ability is represented as being linked across the *verbal-educational*, *kinaesthetic-motor* and *numerical* group factors (p. 47).

We may conclude by a brief reference to the prediction of scientific attainment. The results obtained in several of the researches quoted in the preceding paragraphs are not very convincing, but two earlier workers, Earle³ and Zyve,⁴ claim to have found tests which predict attainment in science efficiently. Earle showed that reasoning problems and analogies in verbal and numerical material correlated most highly with achievement in science. Zyve used various tests including those of reasoning, accuracy of observation, clarity of definition, and experimental bent and obtained maximum predictions ranging between $\cdot 74$ and $\cdot 95$ for different branches of science, but in a more recent comparison of the material with other published aptitude tests⁵ correlations ranging between $\cdot 33$ and $\cdot 57$ were obtained.

It seems, therefore, from the generally low individual correlations and the maximum correlations which rarely exceed $\cdot 6$ that, apart from general-verbal intellectual ability and reasoning, we have yet to discover significant psychological determiners of scientific aptitudes. What then has the psychology of individual differences in intellectual power to offer to the science teacher? As far as his evidence shows, there is no strongly delineated group factor of science ability. Pupils possessing high general ability with some *interest* in practical and scientific matters would seem to make the potential scientists. This is something that must be obvious from school work, but there always is the temptation to hypostatise achievement in one field into a special ability. The psychologist has no definitely conclusive evidence of such an ability.

¹ WARD, A. E. H.: *A Study of Practical Tests in Physics at Sixth Form Level*, M.A., London, 1954.

² VERNON, P. E.: *Structure of Human Ability*.

³ EARLE, F. M.: *Tests of Ability for Secondary School Courses*, X.C.R.E., 1936.

⁴ ZYVE, D. L.: "A Test of Scientific Aptitude, *J. Educ. Psych.*, XVIII, 7, pp. 525-546.

Oct., 1927.

⁵ COOPRIDER, H. A., and LASLETT, H. R.: "Predictive Values of the Stanford Scientific Tests and the Engineering and Physical Science Aptitude Tests," *Educ. and Psych. Measurement*, 8, p. 683, 1948.

IV.—TEACHING AND LEARNING SCIENCE.

Psychological theory offers two basic explanations of learning and both may have particular significance for the teacher of science.¹ First, we have the 'gestalt' view which assumes an *active* learner busily giving meaning to all his experiences. This he does by the law of the 'good' gestalt. He organises his experiences to give as 'good' a meaning as the conditions allow. So a pupil offers the explanation that dissolved air makes an immersed body lighter in the traditional experiment on Archimedes principle. This 'insight' is as 'good' as the conditions permit and the teacher's rôle is to improve upon it by discussion and questioning incorporating similar sequences to those suggested by Duncker. The gestalt theory of 'insight,' 'structure' and active organisation is peculiarly appropriate for science teaching since it is essentially the way by which science has developed. Galileo's false analogy of the column of water breaking under its own weight, Torricelli's correct insight of the supporting air were both 'good' gestalts in the sense that they were the best available under the circumstances, but Galileo's blocked progress, Torricelli's opened it up. Similarly, the geocentric theory of the universe and the phlogiston theory of combustion were 'good' gestalts. Does all this mean that the pupil's activity is to go on undirected? It does not. The teacher's task is to start from the insights of the pupils and direct them appropriately to give them correct meaning—actively widening the pupil's experience. It is only under these conditions that transfer of training from science—through its disciplined method—will take place, as Meredith² has shown. By involving the pupil in this active way, the problem of motivation is also solved. A recent interesting study on transfer in science was carried out by Whellock.³ He attempted to find out how far a knowledge of and right attitude towards scientific method was gained by sixth form pupils and Military College students. His items consisted each of a scientific statement followed by a set of five interpretations, conclusions, opinions or actions and in the second part of the test on "clear thinking," he reproduced part of Well's famous Essay on Dew. A sequence of inferential questions was then based on this material. His results were quite promising, for after removing differences due to intelligence, he showed a significant relationship between the scores on method and attitude and scientific background.

There is a second and lower theory of learning⁴ which utilises only observed behaviour and minimises the spontaneous activity and personal experience underlying the 'gestalt' theory. This theory is applicable in school work at the first fringe of new learning. It also renders an adequate account of that learning which is little more than a convention between teacher and taught as in the learning of symbols in chemistry and in the conventions of chemical equations (even here, however, the pupil's knowledge of Latin and Mathematics may rapidly help to give 'meaning' to this kind of learning). This theory of learning might also apply to the learning of biological terminology. Although behaviourism is limited as an explanation of human learning, it can give a plausible account of transfer of training, by means of "stimulus generalisation" and it does emphasise the need for drill work and finally, it points to the need for motivating the learner. Hull's 'reinforcement' of a particular learning only takes place if the learner is adequately motivated. We tend to assume too

¹ HILGARD, E. R.: *Theories of Learning*, Chaps. 7 and 4.

² MEREDITH, G. P.: *The Forum of Education*, Vol. v, 1927.

³ WHELLOCK, R. B.: *An Inquiry into how far Scientific Method is Gained from Science Education*, M.A., London, 1952.

⁴ HULL, C. L.: *A Behaviour System*, Yale Univ. Press, 1952

readily in school work that motivation is taking place. Of course, in science teaching "natural phenomena" provoke most pupils into learning, but there are times when even the science teacher has to supply the motivation. We may take as a final illustration of the interplay of these two theories the well debated theme on the relative merits of lecture demonstration versus laboratory work. When long exercises have to be carried out the former method is preferable, but, if the exercises are short and the pupils can check and evaluate their own progress without too much direction from the teacher then laboratory work is to be preferred. Lastly, the rôle of the teacher as leader and director of progress also shows that the optimum method of teaching is one that makes the most of the insightful methods advocated by the gestalt theorists. At the same time, however, occasions will arise when a great part of the activity will have to come from the teacher, the pupil accepting the information in a way comparable with behaviourist explanations.

Scientific attitudes and interests are also closely connected with the course of learning science, for they embody or express the higher forms of motivation. There has been quite a lot of research on the relationship of attitude, interest and attainment (in Mrs. Blackwell's three lists there are some eighty-five thesis titles on attitudes and twenty-five on interests) but only a few projects on science topics solely. Presumably, however, several of the studies of secondary school attitudes generally include references to science work.

The relationship between attitudes and attainment is usually given as a correlation coefficient. The most commonly quoted figures are not high (about .25 to .30) and the method of measuring attitude is usually that of Thurstone, Likert or Guttman or variations. One is frequently left wondering whether the concept of attitude so defined and measured may not be a little artificial and superficial and whether more could be achieved by using fewer statements and a 'deeper' probe.

'Interest' in science has been measured by questionnaire, vocabulary and information methods and one study¹ has demonstrated that 'interest' forms a clearly marked group factor in a mixed battery of attainment, ability and interest tests. Furthermore, the information method of measuring interest proved a useful predictor of further progress in school science.

From quite another field the work by Berlyne² on perceptual and epistemic curiosity seems also to be relevant to research on scientific interests and attitudes. He treats these concepts on Hullian lines and his ideas are suggestive in a field where "natural phenomena provoke" the pupil into learning.

V.—SUMMARY.

1.—An attempt has been made to show where educational psychology is relevant to science teaching and in order to do this three topics in science education have been discussed in the light of corresponding psychological theory and results. These were the topics of scientific thinking, scientific aptitude and the teaching and learning of science.

2.—It was suggested that the work of the gestalt psychologists, particularly Duncker and Wertheimer, on the nature of problem solving and creative

¹ ANGUS, L.: *A Comparative Study of the Methods of Interest Measurement in Science and its relation to Ability Achievement*, M.Ed., Durham, 1949.

² BERLYNE, D. E.: *A Theory of Human Curiosity*; *Brit. J. Psy.*, xlv. 3, p. 180, Aug., 1954.

thinking and the writings of Piaget could be helpful to the science teacher in connection with promoting scientific thinking in his pupils.

3.—The findings of factorial analysis do not strongly support the existence of a group factor for science. This appears to be a complex of general intelligence, reasoning, spatial number, and other factors.

4.—Lastly, the bearing of 'gestalt' and 'behaviourist' learning theory on science teaching was briefly outlined. Learning in the science room and laboratory is most appropriately viewed as insightful learning, but a limited place can be found for reinforcement theory. The relevance of attitudes and interests in science was also indicated with a suggestion that perhaps 'deeper' methods of assessing these forces might be tried.

CO-EDUCATION: THE VERDICT OF EXPERIENCE

By R. R. DALE

(Lecturer in Education, University College of Swansea.)

I.—Introduction. II.—Review of Research: (i) Moreton; (ii) W. E. Davies; (iii) Swedish enquiries; (iv) Clark. III.—The writer's investigation: (i) method: questionnaire or attitude scale?; (ii) procedure; (iii) percentage of returns; (iv) results. IV.—Discussion and evaluation. V.—Summary. VI.—Appendix.

I.—INTRODUCTION.

THIS article presents one aspect of a many-sided attack on the problem of the relative value of co-educational and single-sex day grammar schools. It examines the verdict of those who have had practical experience in such schools and assesses in particular, the opinion of those who have taught in both types of school. It is an attempt to provide evidence in a realm where supposition and prejudice have hitherto reigned almost supreme.

II.—REVIEW OF RESEARCH.

In the thirties Valentine¹ did some preliminary work among graduates in training for teaching asking them as old scholars of mixed and single-sex schools to judge the relative value of these schools. But first in this country to investigate the opinion of teachers themselves was Moreton who published, in 1946, a summary of his 1939 Ph.D. thesis.² I myself began in 1946, administering my first questionnaire before seeing Moreton's work, and continuing afterwards as it was clear that my own enquiry would be supplementary as well as possibly confirmatory.³ The third researcher in this field is W. E. Davies, whose work is surprisingly similar to my own. He distributed his questionnaire in 1948 and presented his Ph.D. thesis in 1950.⁴

(i) Moreton's Research.

Moreton, using an attitude scale with a population of teachers from 'secondary schools of all types,' which then did not include secondary modern schools, found a 'great preponderance of opinion in favour of co-education,' with 'surprisingly little difference between the attitudes of the various classes of teachers, male and female, married and single, whether educated in mixed sex or single sex schools.' His final table was:

¹ C. W. VALENTINE: *Psychology and its Bearing on Education*, pp. 118, 573. (London, 1950.)

² MORETON, F. E.: "Co-education—A Statistical Enquiry into the Attitude of Teachers towards Co-education, and a Comparative Study of the Emotional Development of Children trained in Co-educational and other Institutions," Ph.D. Thesis, University of London, Institute of Education, 1939 (summarised in *The Brit. J. Educ. Psych.*, June, 1946).

³ Preliminary articles were published in the *Times Educ. Supp.*, Aug. 28th, 1948, and the *A.M.A.*, Sept., 1949. (Accepted Aug., 1948.)

⁴ W. EMRYS DAVIES: "A Study of the Attitudes of Secondary School Teachers towards Co-education, Differentiation of Curriculum and Sex Teaching," Ph.D. Thesis, London, 1950.

TABLE I

SECONDARY SCHOOL TEACHERS. AVERAGE SCORES IN FAVOUR OF CO-EDUCATION.
SCALE RANGE—PLUS 38 TO MINUS 38.

	Married		Single	
	Mixed Educated	Single Sex Educated	Mixed Educated	Single Sex Educated
Men	15	14	13	10·5
Women	19	19	19	19

Moreton indicates the slightly more favourable attitude of women teachers, and the similarity of scores among the women's groups. (The number of married women is only thirty-four.) Neither of these findings are, however, supported by later and more reliable research. Nor does Davies discover any difference caused by marital status (including children), except the more favourable attitude of married women in co-educational schools (significant at the .05 level). On the other hand, the higher score of the co-educated men, compared with those educated in segregated schools is fully confirmed by other workers, as is the finding that teachers 'who have had experience of co-education have a higher opinion of it than teachers in general.'

Moreton's work is a praiseworthy pioneer effort, and his thesis contains much of value as well as his attitude scale section. However, though his conclusion about the preponderance among teachers of opinion in favour of co-education is correct (though somewhat exaggerated), there are several faults. First, there are blemishes in his attitude scale, as he himself acknowledges. Some items are double-barrelled, and supporters of co-education might well wish to endorse one-half of the item and not the other. For example, item 4: 'Boys and girls need quite different school discipline. This cannot be provided in a mixed school,' and item 50: 'It does not matter much from the point of view of discipline whether the school is mixed or separate. Perhaps the separate school is slightly the easier to work in this respect.' This is scored minus 2. Second, he omits from his scale the question of the appointment of women to the headships of co-educational secondary schools. My own research shows that some women teachers feel strongly on this matter, and its omission may have affected the relative scores of men and women. Third, and much more important, is the low percentage of replies. Four thousand forms were sent out, and only 569 replies received. A 14 per cent. return is a most dubious foundation for a dependable superstructure of conclusions. What happened to the other 86 per cent.? Did they throw their forms away because they were too busy or too lazy, or because they disliked co-education? Fortunately, recent work has shown that the former reason was on the whole the major one.

(ii) *Research of W. E. Davies.*

W. E. Davies used an attitude scale to assess opinion among teachers in secondary schools of all types. From this scale he omitted, on the advice of the 'judges,' the question of the appointment of women to the headships of mixed secondary schools. Like Moreton, he found a favourable attitude to co-education but unlike Moreton, found rightly that men are more favourably inclined than women. His work also showed that those teachers who have been educated in a mixed school or have taught in one, or have been trained in a mixed college,

are more favourable to co-education than their opposite numbers from single sex institutions. Head masters are much more favourable than head mistresses. Teachers in mixed grammar schools are more in favour than are teachers in mixed modern schools. The younger generation of women teachers is much more strongly in favour than is the older.

This was a carefully planned statistical enquiry approved by the Consultative Committee on Curricula Appropriate to the Secondary Stage of Education, appointed by the National Union of Teachers. It enjoyed the use of the Union's administrative machinery. It had the facilities for success and deserved to succeed, but success was endangered by the rather low percentage of replies. Of 827 schools selected only 431 consented to participate, and only 396 sent in replies; 4,851 forms were sent out to the 431 schools, and only 2,826 were returned. Even if we ignore the non-participating schools and take as our criterion the proportion of forms returned to those sent out, this percentage is no higher than 58. And though we might legitimately deduct the 360 or so forms sent out to schools which failed to return them, the percentage rises to no more than 63. This is incomparably better than Moreton's, but critics could still claim that the 37 per cent. who failed to reply might have been mainly opponents of co-education. My own enquiry, however, shows that this was not likely, and that those who did not reply were probably mainly those who, whatever their views, could not be bothered, or who mislaid their forms—not difficult in crowded staff rooms.

Another effect of the non-return of forms was the distortion, in several respects, of the scientific sample which Davies had so carefully prepared. While the surprising thing is the correctness of most proportions between the various groups, such as between men and women in the several types of schools, there is a heavy over-weighting of technical schools and some over-weighting of modern schools, in both cases as compared with grammar schools. Within the grammar school group itself, the proportion of teachers from mixed schools is rather high.¹ Yet these discrepancies affect the results only slightly, partly because Davies selected for some purposes a standard sample of 1,000, and partly because he usually kept each sub-group separate from the rest, e.g., Technical Mixed, Technical Single Sex (Men), Technical Single Sex (Women), Co-educational Grammar Schools (Men), etc.

(iii) *Swedish Enquiries.*²

In 1919 a Commission was appointed to examine Swedish education, including especially the question of co-education. Among other enquiries they asked the opinion of the teachers in mixed schools, who supported on the whole the ultimate findings of the Commission in favour of mixed schools. In 1927 the Swedish Parliament adopted the report.³ Clark gives a translation⁴ of the decisions: '26 lycées will be re-organised to lycées mixtes de l'Etat, in all classes, and 10 only in 'classes supérieures.' 9 lycées de l'Etat will remain for boys only in four large cities. There will be 81 écoles modernes mixtes de l'Etat of which 54 now are écoles communales moyennes (also mixed). Only 5 écoles modernes de l'Etat will remain schools for boys in three large cities. There will also be founded lycées de l'Etat for girls in large cities.'

¹ Percentage of women teachers in mixed grammar schools 39, compared with an estimated true sample of about 29. Percentage of men teachers in mixed grammar schools 46, compared with an estimated true sample of about 33.

² Report of the Commission of 1919. Stockholm, 1923.

³ Quoted in CLARK, G.: "Co-education—An Analysis of the Work and Principles of the Mixed School," M.A. Thesis, Liverpool, 1937.

⁴ Tr. by ANNA B. BERGSTRAND, Stockholm.

In a later enquiry questionnaires were sent to 185 Heads of State municipal co-educational schools.¹ Not all heads replied, nor do all the replies appear to have been classifiable, but 70 replies were very strongly in favour of co-education and wished to see it extended, 30 others were favourable but rather indefinite—writing of co-education as being the only economic possibility and having no great inconveniences, while 20 were very critical.

(iv) *Clark's Enquiry.*²

Clark sent a questionnaire in 1936 to the heads of all mixed secondary grammar schools. He received only 72 replies, but although just under 90 per cent. had taught in segregated schools, 80 per cent. of the heads said mixed schools had marked advantages over segregated schools, and only 8 per cent. were in favour of segregation. Two of the replies are quoted here as they are typical of those received by the writer. 'I was a house master in a public school, and I want no more of it,' and 'For seventeen years I taught in a single-sex school, then I became head of a mixed school and saw the folly of segregation.'

III.—THE WRITER'S RESEARCH.

(i) *Method : Questionnaire or Attitude Scale ?*

Because of the necessity of securing a high percentage of returns and for methodological reasons, an attitude scale approach was discarded in favour of the simple questionnaire—with precautions. In recent years some educationists have frowned on the use of the questionnaire, an attitude which was expressed by McNemar in his well-known article on 'Opinion—Attitude Methodology.'³ The other side of the case, as given in the reasoned replies of Conrad,⁴ Crespi⁵ and others is often ignored. These point out that McNemar made no attempt to present a reasoned discussion of the pros and cons, but used the worst examples of the questionnaire that he could find—with absurdly long or obviously ambiguous single questions. In other words, his attack on the use of the questionnaire became in reality an attack on its misuse—which is a very different matter. Even in the attitude scale technique the 'judges' may arrange the items in an order which is less than perfect, and the scale may omit aspects which are essential to the attitude under examination. We have seen that both Moreton and Davies omitted the problem of the appointment of women to the headships of co-educational schools. Now although this problem might have been at the back of the minds of some women when completing their forms, and therefore have affected their results somewhat, the extent to which its omission affected the validity of the scales is an open question. For some women teachers, especially head mistresses, this problem of promotion might so dwarf all other considerations as to make it impossible to get a true estimate of their opinion on co-education by any known attitude scale. On the other hand, the inclusiveness and directness of the present questionnaire's, 'Are you in favour of co-education in secondary schools?' has advantages, once the word 'co-education' has been carefully defined. Each individual himself weights up the arguments, sub-consciously using his own 'scale values,' without the imposition of arbitrary

¹ "Flickskolan," 1940, *Års Skolulrednings Betänkanden Och Utredningar. Statens Offentliga Utredningar* 1947 : 49. Stockholm, 1947. Tr. by OLE OLDEN, Stavanger.

² *Op. cit.*, p. 262.

³ McNEMAR, QUINN, in *Psych. B.* 43, 289-374, 1946.

⁴ HERBERT S. CONRAD, *Psych. B.*, 43, 6, 1946.

⁵ LEO P. CRESPI, "Opinion—Attitude Methodology and the Polls—A Rejoinder," *Psych. B.*, 43, 6, Nov., 1946.

scale values from outside: he then decides not merely for or against or 'undecided,' but is able to vary the extent of his approval, i.e., whether he favours full co-education (Type A),¹ co-educational schools with single-sex classes (Type B), or Dual Schools (Type C). If he has further reservations he is able to express these in a 'Comment' section. In passing, the possibility is noted that the two sexes might have rather different ideas about the relative importance of some components of an attitude, and if voting on an attitude scale is merely by endorsing items that are agreed with (as in Moreton) and the scale values have been decided by a panel of judges which is predominantly male, the results would be distorted for the female sex. If, however, the voting is of the Likert type, as used by Davies, i.e., on a five-point scale ranging from 'Strongly agree' to 'Strongly disagree,' these discrepancies are reduced. In essence, the voting in the enquiry described here was on a five-point scale.

The balanced view is that there is a place for both questionnaire and attitude scale. The former gives a general conspectus; once this has been obtained other techniques are sometimes desirable. It might also be argued that the direct question in the questionnaire is a check on the complicated machinery of the attitude scale. In other words, if you wish to know the answer to a question, why not ask it? If the question is clear, unambiguous and suited to the population, and the answers anonymous if necessary, the result may be more correct than that obtained through the attitude scale, where faults may be present but difficult to detect, even in the improved versions of Guttman and others.

It remains true, however, that mis-use of the questionnaire is easy, particularly if it employs only one question. Ambiguity and lack of clarity are then especially dangerous. In the present enquiry, co-education was defined and classified, as mentioned above, the field limited to secondary schools, and space provided for comments. A trial run showed the need for small improvements in arrangement, but revealed no ambiguity. The comments section, both in the trial run and in the investigation proper was a useful check on the validity of the polling.²

(ii) *Procedure.*

The forms were sent to each grammar school head who distributed them and collected the completed forms, each in a sealed envelope. To encourage frankness, replies were anonymous. Although the enquiry was conducted in both England and Wales, we are here concerned with one area, the county of Glamorgan.

(iii) *Percentage of Returns.*

Thirty-two schools co-operated; four failed to do so. Of these four one was co-educational (the head was ill), two were for boys and one for girls. As their non-co-operation was due to the unwillingness of the heads to ask their staffs to complete the forms, we are justified in ignoring them when calculating the percentage of returns. There were 550 replies from 665 staff available, a return of almost 83 per cent. Several factors combine to make this percentage even more representative than it appears. For example, the percentage depended largely on the efficiency of the head in distributing and collecting the forms; any failure on the head's part would not necessarily cause bias in the sample. One head, apologising for a 50 per cent. return, explained that he was retiring

¹ See questionnaire in the Appendix.

² The incorporation of an analysis of these comments would have made this article too long; they will be presented separately.

in a few days. Of the four schools which sent in returns of 50 per cent. or less, two were co-educational, one was for boys, and one for girls. The remaining twenty-eight had a return of completed forms of 88 per cent. When we consider the busy life led by teachers, the prevalence of objections to form-filling, the possible non-notification of staff absentees, and mislaid forms—especially where the completed ones were not collected for weeks, we must admit that the sample is representative, especially as the percentages for co-educational and single-sex schools are fairly similar : Co-educational 85 per cent., boys' schools 86 per cent., and girls' schools 77 per cent. (using the very conservative overall average of 83 per cent.).

(iv) *Results.*

The results are shown in Table II. In co-educational schools 84 per cent. of the replies were in favour of full co-education (Type A) and 7 per cent. in favour of co-education with single-sex classes (Type B); only 4 per cent. were undecided. In boys' schools 51 per cent. favoured Type A, 10 per cent. Type B, with 6 per cent. undecided. In girls' schools 41·5 per cent. were for Type A, 13 per cent. for Type B, and 11 per cent. undecided. Whereas an overwhelming majority of the staffs of the co-educational schools preferred their own system, a majority of the staffs of the single sex schools voted against their own system and preferred co-education. In some girls' schools the bitter opposition of the head mistress to co-education may have deterred some assistants from making a true return, in spite of the attempt to preserve anonymity. On the other hand, co-operated willingly in spite of their opinions.

TABLE II
SUMMARY OF REPLIES FROM ALL TEACHERS.
ANALYSIS BY TYPE OF GRAMMAR SCHOOL.

Replies from Staff of :	* Staff Avail- able	Replies	Staff in Favour of :					A + B	S. Sex + C
			Co-educational		Dual Schools C	Undecided	Single- Sex Schools		
			A	B					
		% Staff	% Replies	%	%	%	%		
Co-educational Schools	c 252	215 85·3	181 84·2	15 7·0	5 2·3	8 3·8	6 2·8	196	11
Boys' Schools	c 197	169 85·8	86 50·9	17 10·1	19 11·2	10 5·9	37 21·9	103	56
Girls' Schools	c 216	166 76·9	69 41·5	21 12·7	27 16·3	19 11·4	30 18·1	90	57
TOTALS	665	550 82·7	336 61·1	53 9·6	51 9·3	37 6·7	73 13·3	389	124

* Figures usually supplied by Heads : excluding absentees and most part-time staff. A few assistants of only several months' experience were excluded by the Heads. Replies from several women working in boys' schools were excluded because of the unusual type of experience they represented. The replies of those Heads who elected to answer the questionnaire are included in all tables.

This verdict might be criticised because many of these teachers had teaching experience in only one of the two types of school, so that they could not make a valid comparison. Though the reply could be made that some of these teachers had this experience during their education, this does not meet the whole of the objection. Table III therefore tabulates the opinion of those teachers who had taught in both single-sex and co-educational grammar schools.

TABLE III
STAFF WITH SINGLE-SEX AND CO-EDUCATIONAL TEACHING EXPERIENCE.
CLASSIFIED ACCORDING TO EDUCATION OF TEACHER.

Replies from Staffs of	Education of Teacher	Staff in Favour of :				Undecided	Totals
		A Complete Co-education	B Co-education with mostly Separate Classes	C Dual Schools	Single Sex Schools		
1. Boys' Schools.	(a) Single Sex	25	3	3	8	0	39
	(b) Co-ed.	26	3	2	2	0	33
	(c) Both Types	1	0	0	0	0	1
	Totals	52 (71.3%)	6 (8.2%)	5 (6.8%)	10 (13.7%)	0 (0%)	73
2. Girls' Schools.	(a) Single Sex	17	6	8	8	2	41
	(b) Co-ed.	14	4	3	3	0	24
	(c) Both Types	8	1	1	5	1	16
	Totals	39 (48.1%)	11 (13.6%)	12 (14.8%)	16 (19.8%)	3 (3.7%)	81
3. Mixed Schools.	(a) Single Sex	10W 15M 25	2M	2M	2M	0	31
	(b) Co-ed.	14W 30M 44	2M 1W	1W	0M	2M	50
	(c) Both Types	1W	2M	0	0	0	3
	Totals	70 (83.3%)	7 (8.3%)	3 (3.6%)	2 (2.4%)	2 (2.4%)	84
TOTALS		161 (67.6%)	24 (10.1%)	20 (8.4%)	28 (11.8%)	5 (2.1%)	238

N.B.—Out of a total of 238 replies 154 were from single-sex schools and 84 from co-educational. It is worthy of note that half of these teachers were educated in co-educational schools.

In this table the verdict is once again heavily in favour of co-education. If we combine Types A and B as opposed to Single-Sex plus C, we see that in mixed schools A+B is 94 per cent., in boys' schools 79.5 per cent., and in girls' schools, 64 per cent. An interesting point arises if we split the staff from mixed schools in Table III into men and women :

TABLE IV

STAFF WITH SINGLE-SEX AND CO-EDUCATIONAL TEACHING EXPERIENCE NOW TEACHING
IN CO-EDUCATIONAL SCHOOLS.
(In Percentages.)

	In favour of :				
	A Complete Co-education	B	C	Single Sex	Undecided
Men : N=57 ..	78.9	10.5	3.5	3.5	3.5
Women : N=27 ..	92.6	3.7	3.7	0	0

Here (in Table IV) we see that women teachers in mixed schools, who have also taught in girls' schools, appear to be more strongly in favour of co-education than are men teachers in mixed schools who have also taught in boys' schools. As the numbers are very small, let us consider *all* teachers in mixed schools who replied (Table V).

TABLE V

REPLIES FROM TEACHERS IN CO-EDUCATIONAL SCHOOLS.
(In Percentages).

	In favour of :				
	A Complete Co-education	B	C Dual Schools	Single Sex	Undecided
Men : N=131	78.6	9.9	3	3	5.3
Women : N=82 ..	92.7	2.6	1.2	2.4	1.2

The similarity between the position shown in this table and that in Table IV is somewhat surprising, although the population of Table IV is included in Table V. The percentages of men and women in favour of full co-education remain roughly constant, with the women teachers in mixed schools more favourably inclined than the men. This is interesting because the overall picture, combining teachers in mixed with teachers in single-sex grammar schools, shows the men teachers to have a slightly more favourable attitude to co-education than have the women. This is because women teachers in girls' schools are less in favour than are men teachers in boys' schools. In addition, the teachers from girls' schools have more weight in the final table as they are much more numerous than women teachers in mixed schools.

The influence of the education of the teacher is illustrated in Table VI.

TABLE VI

STAFF WHO HAVE TAUGHT IN BOTH SINGLE-SEX AND CO-EDUCATIONAL SCHOOLS.
SUB-CLASSIFIED ACCORDING TO EDUCATION.

Replies from Staffs of	Educated in Single-Sex Schools	Prefer Full Co-education	Educated in Mixed Schools	Prefer Full Co-education
Boys' Schools	39	25	33	26
Girls' Schools	41	17	24	14
Mixed Schools	31	25	50	44
TOTALS	111	67	107	84
Percentages	—	60·4%	—	78·5%

Here we see that 60 per cent. of teachers who have taught in both single sex and co-educational grammar schools, and were educated in single-sex schools, are in favour of co-education in secondary schools, while the percentage rises to 78 for similar teachers who were educated in co-educational schools. This finding is corroborated by the work of W. E. Davies.

Tables VII and VIII examine the relationship between teachers' ages and their opinion on co-education. They show that whereas there is no discernible trend in the mixed schools, where the women teachers especially show a remarkable consistency, both in the boys' and girls' schools there is a decline of approval in successive age groups, except in the 51 to 60 age group of women (though numbers are too small in this case for reliability). Several women who refused to state their age, and who were undecided in their views or opposed to co-education, may have been from this age group, but their inclusion would have reduced the percentage in favour only to 48. The break in the decline may have been due to chance, or to the passing from the troubled forties to the calmer and

TABLE VII

AGE DISTRIBUTION TABLE.
WOMEN TEACHERS (GLAMORGAN).

Age	Mixed Schools					Girls' Schools					
	Teachers' Preference					Teachers' Preference					
	A	B	C	Und.	Single Sex	A	B	C	Und.	Single Sex	% A+B
21-30	29	2	0	1	1	33	14	4	2	6	80
31-40	22	0	0	0	0	16	6	7	5	5	56·4
41-50	11	0	1	0	0	11	0	10	8	15	25
51-60	14	0	0	0	1	9	1	5	0	3	(55·6)
TOTALS	76	2	1	1	2	69	21	26	15	29	

NOTE.—Percentage A+B is the percentage of A+B to all replies from the same age group in girls' schools. The schools throughout this enquiry are secondary grammar schools. In the Girls' Schools five teachers withheld their ages and could not be classified; four of them were 'undecided' and one in favour of single-sex schools.

more philosophical fifties. The general trend, however, may be partly due to factors other than age, e.g., the lingering Victorian tradition of the separation of the sexes, and selective emigration from girls' schools. The trend is confirmed by Davies, who, however, did not separate out the teachers in mixed from the teachers in single-sex schools. In his case the differences between the older and younger generations of men teachers, taking mixed and single-sex schools together, was not quite statistically significant.

TABLE VIII
AGE DISTRIBUTION TABLE.
MEN TEACHERS (GLAMORGAN).

Age	Mixed Schools						Boys' Schools					
	Teachers' Preference.						Teachers' Preference.					
	A	B	C	U	Single Sex	% A+B	A	B	C	U	Single Sex	% A+B
21-30	11	1	2	2	0	75	15	4	2	1	2	85
31-40	39	5	1	4	1	88	34	6	9	4	12	62
41-50	35	5	1	0	2	93	25	3	5	4	11	58
51-60	16	1	0	1	1	90	11	2	6	1	8	46
Over 60	2	1	0	0	0	—	2	0	0	0	1	—
TOTALS	103	13	4	7	4	—	87	15	22	10	34	—

IV.—DISCUSSION AND EVALUATION.

The preferences which have been cast are not merely theoretical opinions on a subject of which the population has little knowledge, but judgments of practising teachers, based at worst on experience of only one of the two types of school under discussion, at best on experience of both types. On the one hand the overwhelming majority of both men and women teachers in maintained co-educational grammar schools in Glamorgan believe that full co-education (Type A) provides a better education for day pupils at the secondary stage than does segregation. On the other hand, teachers in the single-sex grammar schools do not believe that their own segregation is the better policy; men and to a lesser extent women teachers in these schools also prefer co-education. When we separate out those teachers who have taught in both single-sex and co-educational grammar schools, we find that their judgment is much more heavily in favour of co-education than is the judgment of those whose experience is limited to single-sex schools. It cannot be gainsaid that the great majority of those who have had intimate experience of co-educational day grammar schools prefer this kind of education. Critics might argue that it is mainly those teachers who like co-education who apply to teach in these schools,¹ but even if this were conceded, the argument proves to be double-edged when applied to single-sex schools, for teachers in these schools also prefer co-education. It is clear that criticism of co-education at the secondary stage is appreciably based upon ignorance, that is,

¹ In the boys' schools 47 per cent. of the teachers were co-educated; in the girls' schools 32 per cent.; in the mixed schools 63 per cent. of the men and 61 per cent. of the women.

upon lack of practical experience—particularly teaching experience—in these schools.

To what extent are these findings valid elsewhere? An extension of the enquiry to other parts of Wales has reproduced the general character of the results, though the material has not yet been prepared for publication. In Monmouthshire and England, work is still in progress, but again the results seem to be falling into a roughly similar pattern. Davies found opinion in Wales to be very slightly more favourable to co-education than opinion in England, which is broadly speaking a confirmation of the above interim findings. Three factors—other things being equal—would reduce somewhat, in England, the preference for co-education: the smaller proportion of co-educational schools, so that fewer teachers have teaching experience in them, and also fewer teachers are educated in them, and the presence in England of a public school system—the present enquiry and its extensions are concerned with maintained day grammar schools. All the enquiries, however, are unanimous in confirming the great popularity on educational grounds of co-educational secondary schools among the staffs of such schools, no matter where the investigations have been conducted. The same can by no means be said of single-sex secondary schools.

The preference for co-education, however, may be expressed for other than educational reasons—either consciously or unconsciously: teachers may find such schools pleasant places in which to work. Two points are made in reply. First, that the happiness of the staff will be reflected in the education given to the pupils. Second, that the comments made in support of the preference were almost entirely educational in nature; this, of course, is concerned only with conscious motivation. The same arguments apply to preferences expressed for single-sex schools. More light will be thrown on some of these problems by work still in progress.

Preferences on such questions as these are decided not only by experience, but by the beliefs of society in general. The building, the curriculum, the staffing, etc., are an expression of the belief of society in education and of the intensity and nature of that belief. In Victorian times society believed in the segregation of the sexes, so that co-educational schools appeared to be unnatural institutions. With the passing of that Victorian attitude one would expect a swing towards co-education, after the usual time lag. But this change is not necessarily permanent; a reversal of sentiment is possible, though in Britain, it does not appear likely. We should also keep in mind that an important part of society opinion is parental opinion; on this only a little work has as yet been done.¹

The assessing of opinion, even though this opinion be based on experience, cannot of itself be conclusive, though its value must not be under-rated. A minority can be right and a majority wrong. A reformer may even be alone against the rest. But the findings of this enquiry gain in significance when it is realised that the preference for co-education is registered in spite of school loyalty and in face of a tradition and an educational press which in this country favours single-sex education, and that the change has taken place in spite of the superior prestige of the single-sex public school system and of the old-established single-sex city grammar and high schools.

The problem of the relative values of co-education and single-sex education has many aspects, and much more investigation is needed before a final judgment can be made.

¹ Cf. the thesis of W. E. DAVIES.

V.—SUMMARY OF CONCLUSIONS.

1.—About 73 per cent. of men and 68 per cent. of women teachers in maintained day secondary grammar schools in Glamorgan prefer co-educational to single-sex day grammar schools. About 6 per cent. of the men and 8 per cent. of the women are undecided. If these undecideds are omitted, the percentages in favour of co-education rise to 75 for men and 74 for women.

2.—The women in co-educational schools are rather more favourably inclined than are the men in these schools (93 per cent. of women in favour of Type A and 79 per cent. of men).

3.—Teachers educated in co-educational secondary schools are more favourable than are teachers educated in single-sex schools, the percentages (for Type A) being 78.5 of all replies and 60.

4.—Those teaching in co-educational grammar schools show a much more decided preference for co-education (91 per cent. of all replies were for Types A+B) than do those teaching in single-sex grammar schools (61 per cent. boys' and 54 per cent. girls'). In mixed schools 4 per cent. are undecided, in boys' 6 per cent., and in girls' 11 per cent.

5.—The experience of those teachers who have taught in both co-educational and single-sex grammar schools leads them to prefer co-education; 84 per cent. of the men and 70 per cent. of the women prefer schools of Type A or B, and of these a great majority prefer Type A.

6.—In the single-sex schools the younger generation of both men and women teachers show a more favourable attitude than do the older generation.

7.—The above findings are based on an 83 per cent. return. For reasons already given, this is even more representative than the figure suggests.

8.—Opposition to co-education springs appreciably from ignorance, i.e., from lack of practical experience in these schools.

9.—The writer's further investigations in England and Wales (incomplete), together with the findings of W. E. Davies, indicate that these findings are probably valid elsewhere in England and Wales, with some slight reduction in England of preference for co-education.

10.—This research has explored only one aspect of the problem; other aspects are under investigation.

VI.—APPENDIX.

CO-EDUCATION—AN ENQUIRY

FORM A.

The enquiry set out below is part of an investigation into the question of co-education. It is hoped that all members of school staffs, of university departments of education and training colleges, to whom this form is sent, will assist the enquiry by answering the questions and adding their observations. All replies will be strictly confidential; members need not sign the statement unless they expressly wish to do so.

In order to make the enquiry sufficiently precise, co-education is classified into three types:

Type A—Where boys and girls are educated together in the same classes (except for such subjects as needlework and physical training) and in school society activities.

Type B.—Where boys and girls are in the same school under one head master (or head mistress), and have joint social activities, including school societies, but are separated for all classwork.

Type C.—Where boys and girls of a school are in separate departments, under separate heads, but there is a joint social life, including school society activities.

Are you in favour of co-education in Secondary* Schools? Yes/No/Undecided.

If you are in favour, indicate which of the above types you prefer : Type

Length of teaching experience in co-education secondary schools : Type A years
Type B years
Type C years

Length of teaching experience in single-sex secondary schools years

Were you educated in a co-education secondary school? Yes/No/Part of the time.

Which type? Type

For how long? years

Please state your age years
and sex Male/Female.

If you have time, please state briefly below one or two of the principal reasons for your opinion on co-education.

* This means here all schools which are for pupils in their teens, but not including, e.g., approved schools.

If you have any special reservation about technical schools, please insert this in the 'comments' section.

THE EVIDENCE FOR THE CONCEPT OF INTELLIGENCE

By CYRIL BURT

I.—*The non-statistical evidence*: (1) *observational*; (2) *biological*; (3) *physiological*; (4) *individual psychology*. II.—*The statistical evidence*: (1) *the general factor*; (2) *the factor as cognitive*; (3) *the factor as innate—the hypothesis of multi-factorial inheritance*. III.—*Summary*. IV.—*References*.

I.—THE NON-STATISTICAL EVIDENCE.

Current Criticisms.—The concept of intelligence, and the attempt to measure intelligence by standardized tests, have of late furnished a target for vigorous attack. The objections urged are partly practical and partly theoretical. Yet few of the critics show a clear or correct understanding of what the term really designates or of the reasons that have led to its introduction. Two misconceptions have become widely current.

(i) Those writers who are chiefly interested in the more practical issues, like Dr. Heim and Dr. Blackburn, explain that intelligence "is a popular and relatively unambiguous word," and denotes a quality that "all can recognize, though few can define."¹ It follows that, instead of pinning I.Qs. on to the coat of each child, we should leave any decisions that may be necessary to the intuitive insight of the teacher. Unfortunately, in a vain effort to measure the immeasurable, the modern psychologist "has been induced to restrict the meaning of the term to a vague quantitative abstraction." No two of them, however, agree as to how that abstraction is to be defined. Hence "those who go chasing this *ignis fatuus* get quickly bogged down in mathematical abstruseness." Meanwhile, the layman, so Mr. Richmond assures us, has begun to "sense a certain absurdity in measuring something called 'intelligence' without knowing what that something is or how it is defined."²

(ii) Those who are concerned with the more technical aspects of the subject apparently suppose that the concept was invented by a small band of statistical enthusiasts—Dr. Kirman (13) mentions Spearman, Pearson, and myself—who deduced their theories by primitive factorial procedures that have since been "publicly discredited." The more accurate methods of Thurstone and his American followers, it is said, have since clearly shown that the intellectual achievements of different individuals are the product, not of a single general factor, but of a number of more specialized 'primary abilities.'³ And this at once accounts for the difficulties that beset all attempts to define intelligence. As Captain Kettle observed, when asked why the pictures of the Saghalien sea-serpent showed such incredible differences: "'Spects it's because there's no such crittur'; so each just draws his own fancy."

The Definition of Intelligence.—Now the critics who protest about "the spate of incongruous definitions" usually rest their complaint on the results of the famous Symposium organized some twenty years ago.⁴ The Editor of an

¹ (11), pp. 30f. Cf. also J. BLACKBURN: *Psychology and the Social Pattern* (1945), p. 61

² (18), p. 227. Similar criticisms have also been put forward by Dr. E. G. Chambers, Dr. D. H. Stott, and Dr. C. M. Fleming.

³ For a recent statement of the American view, see A. ANASTASI: *Psychological Testing* (1955), pp. 15, 353f.

⁴ "Symposium on Intelligence and its Measurement," *J. Educ. Psych.*, XII, 1921, pp. 123-147 and 195-216. In framing his question, the Editor specifically asked, not how is intelligence to be defined, but "what do you conceive intelligence to be, and how can it best be measured: should the test material call into play analytical and higher thought processes, or should it deal rather with simple, with associative, or with perceptual processes, etc."

American journal submitted two searching questions about the nature of intelligence to a dozen different psychologists, and received a dozen different replies. But the varying descriptions suggested were not, as Dr. Heim and others have supposed, intended to be 'definitions' in the strict logical sense: they were, in the language of J. S. Mill, merely "attempts to explain the thing," not "attempts to interpret the word." As the editorial letter shows, the purpose of the discussion was primarily a practical one—to determine how intelligence appears to operate, with a view to ascertaining "what material may most profitably be used in constructing tests." But that is quite a separate question, and except incidentally will not concern us here. Nor shall I discuss the validity of mental measurement or the practical value of the I.Q.¹—problems that are continually confused with the fundamental issue. The questions I now want to settle are prior to all these, namely, (i) how precisely should the term be defined, and (ii) what evidence is there for believing that something really exists corresponding to the definition proposed? However, instead of taking the term for granted and hunting round for a plausible formula, as is most frequently done, a sound scientific procedure requires us to start with the relevant facts. Let us, therefore, take the second of our two questions first.

History of the Concept.—Many of the criticisms to which I have alluded spring largely from a manifest ignorance as to how the concept originated. A rapid glance at the literature is, therefore, needed first of all.² As a brief historical review will show, long before the advent of statistical analysis, several converging lines of evidence had already drawn attention to an important property of the mind, for which some special name seemed desirable. How its nature was envisaged can best be gathered by recalling the actual statements of leading authorities in each field.

(1) *Observational.*

The earliest attempts to analyse and classify the activities of the mind were based partly on the observation of various types of person in everyday life and partly on introspection. Plato, to whom we owe the basic distinctions, draws a clear contrast between 'nature' and 'nurture' (*φύσις* and *τροφή*); and then distinguishes three parts or aspects of the soul—*τὴ λογιστικόν*, *ἰπιθυμικόν*, *θυμικόν* (*Republic*, 435A). The modern terms—intellectual, emotional, and moral, cognition, affection, and conation—suggest rough but somewhat inexact equivalents for these untranslatable expressions. In a celebrated passage (*Phaedrus*, 253D) he sketches a picturesque analogy which conveys a better notion of the fundamental difference: the first component he compares to a charioteer who holds the reins, and the other two to a pair of horses who draw the vehicle: the former guides, the latter supply the power; the former is the *cybernetic* element, the latter the *dynamic*.

Aristotle makes a further contribution of lasting importance. He

¹ For a discussion of these questions I may refer to Professor Vernon's address on 'The Psychology of Intelligence and G' in the current *Bulletin of the Brit. Psychol. Society* (No. 20, pp. 1-14), which I had not seen before this article was written.

² A more detailed account will be found in my "Historical Sketch," which forms the first chapter of the Board of Education Report on *Psychological Tests of Educable Capacity* (2, pp. 1-61) and in a recent Galton Lecture on "The Meaning and Assessment of Intelligence" (5). The antecedent evidence, drawn from the four main fields reviewed below, was briefly summarized in my earliest papers on general intelligence (e.g., *J. Exp. Pedag.*, 1911, pp. 96). If the reader refers to that article, he will see that the criticism made by Dr. Maberley, and repeated in varying terms by several later writers—namely, that I "claimed to deduce the general factor from a statistical analysis of test-data"—quite misrepresents my argument: the statistical analysis was intended merely to confirm a hypothesis reached on far more concrete grounds.

contrasts the actual or concrete activity with the hypothetical capacity¹ on which it depends (*δύναμις*), and thus introduces the idea of an 'ability.' Plato's threefold classification he reduces to a twofold. For him the main distinction is between what he calls the 'dianoetic' (cognitive or intellectual) capacities of the mind and the 'orectic' (emotional and moral).¹ Finally, Cicero, in an endeavour to supply a Latin terminology for Greek philosophy, translates *δύναμις* by *facultas*, and *ὀρεξις* by *appetitus* or sometimes *conatus*; while to designate *διανοία* he coins a new word, rendering the Greek term almost literally by the compound '*intellegentia*.'

Here then we have the origin of both the concept and the term. So far from being a 'word of popular speech,' whose meaning has been restricted and distorted by the modern psychologist, intelligence is a highly technical expression invented to denote a highly technical abstraction. From Aristotle and Cicero it descended to the mediaeval schoolmen; and the scholastic theories in turn became elaborated into the cut-and-dried schemes of the faculty psychologists and their phrenological followers.

(2) *Biological.*

As Guilford has reminded us, the modern notion of "intelligence as a unitary entity" was "a gift to psychology from biology through the instrumentality of Herbert Spencer." Following Aristotle and the later Scottish school, Spencer recognizes two main aspects of mental life—the cognitive and the affective. All cognition (he explains) involves both an analytic or discriminative and a synthetic or integrative process; and its essential function is to enable the organism to adjust itself more effectively to a complex and ever-changing environment. During the evolution of the animal kingdom, and during the growth of the individual child, the fundamental capacity of cognition "progressively differentiates into a hierarchy of more specialized abilities"—sprouts into boughs, branches, and twigs. To designate the basic characteristic he revives the term 'intelligence.'²

Evidence favouring Spencer's somewhat speculative theories was adduced by Romanes, Lloyd Morgan, and other pioneers of comparative psychology; and his views on intelligence were accepted, not only by British biologists like Darwin, but also by continental writers, like Binet and Claparède.³ Certainly, Mendel's earliest disciples maintained that the doctrine of unit-characters was utterly irreconcilable with the inheritability of a graded trait, such as intelligence (cf. 6, pp. 333f.); but, as we shall see in a moment, the later developments of the Mendelian hypothesis not only permit it, but actually suggest it.

(3) *Physiological.*

The clinical work of Hughlings Jackson, the experimental investigations of Sherrington, and the microscopical studies of the brain carried out by Campbell, Brodmann, and others, have done much to confirm Spencer's theory of a

¹ DE ANIMA, II, 3, 414a, 31. *Eth. Nic.*, I, 13, 18, 1102b, 30. The usual rendering 'power' must not be taken to imply causal agency: Aristotle is simply describing what Professor Broad has called a 'dispositional property.'

² H. SPENCER: *Principles of Psychology* (1870). I have summarized Spencer's views more fully in a recent article ("The Differentiation of Intellectual Ability," this *Journal*, XXIV, 1954, pp. 76f).

³ Cf. C. DARWIN: *The Descent of Man* (1888), I, pp. 101f.; G. J. ROMANES: *Animal Intelligence* (1890); and LLOYD MORGAN: *Animal Life and Intelligence* (1796).

'hierarchy of neural functions,'¹ with a basic type of activity developing by fairly definite stages into higher and more specialized forms. In particular, the examination of the cortex, both in mental defectives and in normal persons, suggests that the quality of the nervous tissue in any given individual tends to be predominantly the same throughout. Defectives, for example, exhibit a "general cerebral immaturity"; their nerve-cells tend to be "visibly deficient in number, branching, and regularity of arrangement in every part of the cortex."² After all, as Sherrington himself points out, much the same is true of almost every tissue of which the human frame is composed—of a man's skin, bones, hair, or muscles: each is of the same general character all over the body, although minor local variations are usually discernible. In the adult human brain marked differences in the architecture of different areas and of different cell-layers are perceptible under the microscope; but these specializations appear and develop progressively during the early months of infant life. And, of course, such differentiation is precisely what the Spencerian theory would entail.

The experimental study of the brain leads to the same conclusion. The intact brain acts always as a whole. No part of the brain functions in total isolation from the rest, as the older champions of cortical localization originally assumed. The activity, in Sherrington's phrase, is "patterned not indifferently diffuse"; but the patterning itself "involves and implies integration." Lashley's³ conclusions about the 'mass action' of the brain seem to lend further corroboration to that view; and, as several writers have suggested, this 'mass-action' might well be identified with *g*.⁴

The evidence of neurology, therefore, itself suggests something very like a theory of general ability, which gradually differentiates into more specific functions, though we must beware of picturing such functions as separate 'faculties' located in certain centres or compartments of the brain, after the fashion of the older phrenologists and of several recent writers on so-called 'physiological' or 'medical' psychology.

(4) *Individual Psychology.*

All these earlier writers were interested primarily in the working of the mind as such, that is to say, in problems of *general* psychology. The first to apply scientific methods to the problems of *individual* psychology was Galton. Darwin and Spencer had maintained that the basic capacities of the human mind were hereditary, transmitted as part of our common racial endowment. Galton went farther and maintained that individual differences in these capacities were also innate. As a result of his investigations into 'hereditary genius,' he was led to discard the traditional explanation in terms of faculties and types, and to substitute a classification in terms of 'general ability' and 'special aptitudes':

¹ The phrase is Sherrington's. Cf. C. S. SHERRINGTON: *Integrative Action of the Nervous System* (1906), pp. 314f; HUGHLINGS JACKSON: *Brain* (1899), XXII, pp. 621f.; M. DE CRINIS, "Die Entwicklung der Grosshirnrinde in ihren Beziehungen zur intellektuellen Ausreifung des Kindes," *Wiener Klinische Wochenschrift*, 1932, XLV, pp. 1163f.; J. L. CONEL: *The Post-Natal Development of the Human Cerebral Cortex* (1941).

² J. S. BOLTON: *The Brain in Health and Disease* (1914).

³ K. S. LASHLEY: *Brain Mechanisms and Intelligence* (1929). The experiments of Lashley and his colleagues consisted in training animals to perform definite tasks, and then removing parts of their brains: the animals were then re-tested, and in some instances re-trained. The main conclusion was that ability to learn depends, not so much on the nature or location of the tissue remaining, but upon its amount.

⁴ This identification is suggested by SHERRINGTON (*Man on His Nature*, 1940, p. 288). It should be added that the details of Lashley's conclusions are not entirely free from criticism; but here we are concerned only with the major principle.

of the two he considered that general ability was "by far the most powerful".¹ The differences between individuals formed, so he believed, not a set of distinct and discontinuous classes, as the type-theory assumed, but a series of continuously varying gradations, distributed more or less in accordance with the normal curve, i.e., much like differences in head-length, arm-length, or stature (10, pp. 23f., 35f.).

The Definition Implied. These converging lines of inquiry, therefore, furnished strong presumptive evidence for a mental trait of fundamental importance defined by three verifiable attributes: first, it is a general quality; it enters into every form of mental activity; secondly, it is (in a broad sense of the word) an intellectual quality—that is, it characterizes the cognitive rather than the affective or conative aspects of conscious behaviour; thirdly, it is inherited or at least innate; differences in its strength or amount are due to differences in the individual's genetic constitution. We thus arrive at the concept of an *innate, general, cognitive ability*. We cannot, however, keep repeating a cumbersome phrase of twelve syllables every time we wish to mention it. And, since a name that suggests its own meaning seems preferable to a brand-new esoteric symbol, what better label can be found than the traditional term 'intelligence'?²

Here then is a clearly formulated hypothesis, the outcome of centuries of shrewd observation and plausible conjecture—a psychological hypothesis fully in accord with the findings of the biologist and neurologist. Nevertheless, each of the three propositions that I have just laid down has been vigorously challenged; and each has started off a protracted controversy that still remains unresolved.

At this point, therefore, the need for *ad hoc* inquiries based on rigorous statistical analysis becomes obvious. It is the function of statistical procedures to decide between alternative hypotheses by testing their verifiable corollaries. The claim of the factorist is not, as his critics so often imagine, to 'discover' mental abilities, running round with a cry of 'Eureka' whenever he has extracted a fresh factor: his object is merely to confirm or refute certain hypothetical concepts or components that have been tentatively reached on more concrete grounds. Let us then take each of the three foregoing propositions in turn, and consider what evidence, if any, is provided by these more cogent techniques.

¹ Many contemporary writers, particularly in the field of education, attribute the antithesis between 'general' and 'special' abilities to Spearman, and identify it with the contrast between what he called *g* and *s*. Spearman himself, however, frankly admits that his own theories were prompted by those of Galton and Spencer. However, in his earlier discredited faculties of the older school—the only 'specific' capacities that he recognized were those 'specific' to each particular test (cf. *Amer. J. Psych.*, XV, 1904, pp. 74f, and 206f, and 20, pp. 6f.).

² In educational psychology the popularity of the term is due to the work of Alfred Binet, himself an avowed follower of Spencer and Galton. Like Galton, Binet firmly believed in the existence of a 'general ability,' and repeatedly distinguished it from what he called 'partial aptitudes.' This ability, he says, enters into "nearly all the phenomena with which the experimental psychologist has previously concerned himself—sensation, perception, memory, as well as reasoning," i.e., it is essentially a cognitive capacity. Finally, he explains that his intelligence tests were deliberately constructed to measure innate differences, in contrast to his pedagogical tests which measure acquired attainments (cf. esp., *L'Année Psychologique*, XI, 1905, pp. 191f., 245f.).

Galton himself more frequently spoke of 'general ability.' But at times he used 'intelligence' as a synonym, especially when the context called for the adjective (e.g. 9, p. 336). Those who fear the ambiguities of the more familiar term can use a literal symbol: I have suggested using γ for the hypothetical quality defined as above, and keeping *g* for the empirical measurement, with a subscript to indicate the method of measurement.

The logic of the argument should be carefully noted. In the natural sciences a direct deductive proof is out of the question: the mode of proof must be indirect and inductive. Hence, the conclusions reached can never be certain, but only probable. The critic commonly misses this point. He revels in demonstrating that some alternative interpretation can readily be conceived. But one can always think up alternatives. The verdict must depend on determining and balancing the crucial facts. A probable hypothesis can only be overthrown by showing that its rival is still *more probable*. And equally, of course, the defender of a hypothesis must prove that every alternative that is worth considering is *less probable* than his own.

II.—THE STATISTICAL EVIDENCE.

(1) *The General Factor.*

At the beginning of the century, the problem which chiefly exercised students of individual psychology was, in Bain's phrase, 'the classification of intellectual abilities or powers.' (i) Were there, as the faculty psychologists maintained, a number of specialized abilities, each independent of the rest—observation, practical ability, memory, language, reasoning, and the like? (ii) Or was there, as Ward maintained, "not a congeries of faculties, but only a single subjective activity"—a general capacity for cognition as such? (iii) Were there, as Galton believed, both a general ability and a number of more or less specialized capacities? (iv) Or, finally, might there be, as the earlier associationists and most of the later behaviourists alleged, no discernible structure in the mind at all?

Each hypothesis entailed its own distinctive corollaries; and Galton's technique of correlation offered a ready-made method of checking them. Thus, the obvious plan for attacking such a many-sided issue was to devise and apply suitable tests for the main forms of mental activity, and then calculate the correlations between each test and the rest. If, for example, the orthodox behaviourist is right, and there is "no organized structure in the mind—no ground for classifying mental performances under one or more broad headings, no basis for inferring efficiency in one type of activity from efficiency in another," then we should expect *all the intercorrelations to be zero or at least non-significant*.¹

¹ Spearman, writing of the "momentous investigation by Cattell and Wissler" "the first to apply 'the Galton-Pearson coefficient of correlation' to the results obtained with psychological tests—evidently understands them to have accepted this inference (20, p. 56). Wissler, it is true, says that at first sight the low coefficients suggest that "every act measured by the tests is special and unrelated to every other act" (22, p. 55): but he plainly does not intend this conclusion to be final: he speaks of a "deep conviction that we are otherwise constituted," and points out that certain correlations (e.g. for memory and college grades) are positive and significant. Thorndike also said it was tempting to infer from the data that "there is *nothing whatever* (his italics) common to all mental functions or to any part of them" (*Amer. J. Psych.*, XX, 1909, p. 368): but he, too, quickly abandoned this view. The reasons for the low correlations obtained in these earlier researches are now quite clear: (a) the earlier tests had a low reliability; (b) the functions tested were extremely simple, and the size of the correlation tends to increase with the complexity of the function; (c) the groups tested (students or school classes rather than complete age groups) were already highly selected for general intelligence.

Thomson's sampling theory, though expressed in language similar to that of the 'anti-structural psychologists,' leads to very different corollaries. "The Mind," he says, "has little structure: unlike the body, it is not sub-divided into distinct organs, but forms a comparatively undifferentiated complex of innumerable elements." These he pictures as 'bonds,' i.e., interconnecting neural paths: they have the same character or quality throughout the brain. But, so far from the effects of specific stimuli being limited to specific neural paths (as the earlier opponents of structure assumed), "any sample whatever of these elements can be assembled in the activity called for by a 'test'" (21, pp. 303, 306). Now

If, on the other hand, the mind consists of a number of specialized faculties or abilities, such as 'observation' (assessed by tests of sensory capacity) or 'practical ability' (assessed by tests of motor capacity), then we should expect that all the inter-correlations between the sensory tests would be positive and similarly that all the inter-correlations between the motor tests would be positive; on the other hand, we should expect that *all the cross-correlations between the one group and the other would be approximately zero*. If what Thorndike called 'the theory of natural compensation' held good, then the cross-correlations would actually become negative, since the 'sensory type' would be deficient in the characteristic capacities of the 'motor type' and *vice versa*. Lastly, if there were no specific faculties at all, but only 'a single cognitive activity'—'attention,' as Ward believed, 'sensory discrimination' as Sully maintained—then we should expect the entire table of correlations to exhibit what Spearman called a 'perfect hierarchical order,' or (in the more precise language of the mathematical textbook) to have 'a rank of one'—apart, of course, from minor aberrations due to sampling errors.

The results of the earlier inquiries revealed, almost without exception, *positive and significant correlations between every form of cognitive activity*. This disproves hypotheses (i) and (iv). Further, except when the sample was small and the sampling errors large, there were nearly always *well-marked clusters of augmented correlations confined to similar forms of cognitive activity*, and leaving significant residuals after the general factor was removed.¹ This rules out hypothesis (ii). We are thus left with hypothesis (iii) as the only alternative consistent with the facts. And, accordingly, the unavoidable inference is that *both* a 'general factor' and a number of 'group factors' must be at work.¹

But we are not yet justified in identifying this abstract 'general factor' with anything so concrete as 'general intelligence.' In Spearman's investigations 'general intelligence' is always represented by an *external* criterion, i.e., either by direct assessments for intelligence as popularly understood or (in later researches) by standard tests, selected as furnishing accredited 'reference values.' In my own investigations, the 'general cognitive factor' forms an *internal* criterion, namely, what I called the 'highest common factor' in the battery of tests. And to determine the concrete nature of such a factor, or

this (as Thomson recognizes) is merely another version of the general factor theory: the chief difference is that with Spearman the general factor is identified with something concrete (mental energy); with Thomson it represents something abstract (the fact that the neural elements have the same general character throughout). The corollaries are plain. First, since "the physical body has an obvious structure," the contribution of the general physical factor should be much smaller for correlations between bodily measurements than for correlations between mental: indeed, it was this supposed 'contrast with physical measurements' that led Thomson to promulgate his theory. Secondly, with mental measurements, the correlation table, even if not as completely hierarchical as Spearman believed, ought always to exhibit a 'low rank.' Recent work has falsified both these corollaries. To begin with, in the very table for physical measurements which Thomson cites, the contribution of the general factor is practically the same as for mental measurements (50 per cent. or rather more, *Brit. J. Psych., Stat. Sec.*, II, p. 116); secondly, the application of mental tests to much larger samples shows that the low rank of the tables Thomson has in mind resulted from the small numbers tested, whereas the physical measurements were obtained from 3,000 persons. It may be added that no neurologist would subscribe to the view that a stimulus, whether simple or complex, merely 'sampled' the neural elements: the responses to the simpler stimuli are relatively specific and selective; the response to more complex stimulation essentially involves the integration or organization of the neural elements.

¹ C. BURT: "Experimental Tests of General Intelligence," *Brit. J. Psych.*, III, 1909, pp. 94-177.

rather of the processes that give rise to it, a supplementary investigation is requisite, based on observations or introspections, or on the correlation of the factor measurements with independent gradings.¹

Later investigators, notably Brown, Thomson, and more recently Thurstone, have argued that, if we accept the existence of group factors or 'primary abilities,' we can dispense with the hypothesis of a general factor by assuming that the group factors overlap. But this solution has proved unworkable both in theory and in practice. When the general factor accounts for much more of the variance than any single group factor, or indeed than all the group factors put together, there is no theoretical gain in closing one's eyes to its presence. And in educational practice the rash assumption that the general factor has at length been demolished has done much to sanction the impracticable idea that, in classifying children according to their varying capacities, we need no longer consider their degree of general ability, and have only to allot them to schools of different types according to their special aptitudes; in short, that the examination at eleven plus can best be run on the principle of the caucus-race in Wonderland, where everybody wins and each gets some kind of prize.²

In their more recent writings, most of the opponents of the 'general factor hypothesis' have, more or less openly, withdrawn their opposition. Brown, for example, ultimately acknowledged that "the evidence for a general factor now seems conclusive." Thomson himself has constructed numerous booklets for testing intelligence. And Thurstone has proposed a scheme of 'second order factors' which shall expressly include a 'general factor' and so account for the correlations between the 'first order factors' or 'primary abilities.'³

(2) *The Factor as 'Cognitive.'*

Merely to demonstrate the presence of a general factor common to all cognitive activities does not (as is usually assumed) prove that this factor is specifically cognitive. One might as well argue that, because a general factor can be demonstrated common to all sensory activities, therefore this factor is simply and solely a capacity for sensory discrimination. Impressed by this obvious fallacy, a number of writers went on to argue that in all probability the factor common to mental and scholastic activities was not cognitive but conative. Such an interpretation had a warm appeal for those who cherished the doctrine of intellectual equality. When a pupil lagged behindhand in nearly

¹ Actually teachers' gradings for 'intelligence' (as I showed in my 1909 research) are markedly biased in favour of memory or capacity to learn; and many psychologists (e.g., Colvin) adopted this as a definition of intelligence. Spearman, following Sully and the sensationalist school, originally equated intelligence with 'sensory discrimination,' as the basic form of mental analysis. Ward, Stout, and others inclined to identify it with 'attention' or 'apperception,' i.e., mental or 'neotic' synthesis. This early disagreement about the 'nature of intelligence' is no reason for repudiating the concept: after all, there is little agreement about the 'nature' of gravity; but that is no reason for discarding the principle. And, in point of fact, the conflict can easily be reconciled if we borrow the suggestion of the neurologists and suppose its function to be that of 'integration,' i.e., organization (which involves both analysis and synthesis).

² For a fuller discussion of these practical consequences, see this *Journal*, XIII, p. 136, and XXIV, p. 87.

³ Cf. W. BROWN and W. STEPHENSON: "A Test of the Theory of Two Factors," *Brit. J. Psych.*, XXIII, 1933, pp. 352-370; G. THOMSON, *loc. cit. sup.*; L. L. THURSTONE: *Multiple Factor Analysis*, 1947, pp. 421f. As both Brown and Thomson indicated, their change of front was partly the effect of the change in physiological views regarding cerebral localization (notably the conclusions of Lashley in regard to 'mass action' to which they both refer, and Head's drastic criticisms of the 'cerebral map-makers'). Thurstone and his followers, on the other hand, seem indifferent to biological, physiological, or experimental evidence, and prefer to rely exclusively on statistical analysis.

every subject, the teacher was apt to lay the blame on what Dr. Ballard dubbed the 'general factor of laziness.' Conversely, when a bright child forged ahead in all he undertook, he found himself applauded as a paragon of industry and held up to his fellows as a model of zeal: "genius," said the apostles of the gospel of work, "is just an infinite capacity for taking pains."

This interpretation was elaborated in some detail by Maxwell Garnett, Pearson's brilliant assistant, and one of the ablest champions of the doctrine of a general factor. After re-analysing a good deal of the available data, he came to the conclusion that the factor was after all a factor of Will rather than of Intelligence, and affected moral behaviour quite as much as intellectual success.¹ It was largely as a result of his discussions with Garnett that Spearman eventually dropped his earlier interpretations ('sensory discrimination' in his first paper, 'neural plasticity' in the second) and proposed instead a hypothesis of 'mental energy.'

But a re-analysis of existing data, coupled with a priori arguments, could scarcely suffice to settle the question, either one way or the other. Accordingly, in our later experiments, Mr. Moore and I correlated assessments for intellectual performances with assessments for physical, temperamental, and moral qualities. This time most of the cross-correlations were certainly positive, though never very large: it seemed, in fact, as if there was a small but far more comprehensive general factor—a super-factor, as it were—making for excellence in every direction, while the older and more conspicuous factor for cognitive efficiency now appeared simply as a broad group factor, confined to cognitive activities alone: in short, the so-called 'general cognitive factor' turned out to be merely one of the largest of a number of 'group factors' varying in extent and size (2, p. 19). At the same time, another broad group factor emerged underlying the temperamental and moral assessments: this was obviously identifiable with what we had previously called 'the general factor for emotionality.' No sharp division appeared, separating affective characteristics from conative. And the so-called cognitive factor was found to be quite as prominent in tests of practical efficiency as in tests of intellectual activity in the narrower sense.

In the light of this further evidence, Garnett's arguments no longer required us to surrender the idea of a cognitive factor. But it certainly seemed necessary to revise the implications conveyed by the word cognition. The basic contrast seems to lie, not so much between cognitive processes and non-cognitive (i.e., affective or conative) in the old introspective sense of those terms, but rather between the capacity for adapting, guiding, or directing mental activities, by means of discriminative and integrative processes, and the capacity for responding promptly, actively, and energetically. Some such distinction was implicit in Spencer's antithesis between mental mechanism and mental force (or, as the Americans preferred to call it, 'drive'). It was, indeed, the distinction originally laid down by Plato. And, in the absence of more appropriate English names, it is tempting to borrow from the Greek, and speak of a general 'cybernetic'² factor and a general 'dynamic' factor.

¹ J. C. M. GARNETT: *Proc. Roy. Soc., A*, XCVI (1919), pp. 102f. Cf. also *id.*, *Education and Citizenship*, 1921, pp. 476f. It should be noted that in all his writings Garnett, one of the noblest quakers of his day, invariably placed ethical considerations first.

² I.e., a factor for guiding or controlling: see above, Sect. I (1). On the basis of purely observational and experimental work with children, Professor Piaget seems to have reached a very similar interpretation of the traditional antithesis between cognitive and affective processes: cf. *The Psychology of Intelligence*, 1950, pp. 4f.

(3) *The Factor as 'Innate.'*

The evidence we have so far considered seems fully to vindicate the notion of a 'general cognitive factor.' However, during the last fifteen years or so, the most frequent object of attack has been the assertion that this general factor is largely, if not wholly, innate. This line of criticism is partly an after-effect of the doctrines popularized by the behaviourist school, which dominated psychology for so long in the United States. Educational writers in this country still quote Watson's well-known pronouncements: "We no longer believe in inherited capacities . . . All have equal chances at birth."¹ Watson, however, overstated his case. A doctrine of perfect equality in regard to innate mental traits would fly in the face of all biological experience: throughout the animal kingdom, except where the characteristic is absolutely essential to life, innate differences between individuals are the invariable rule.

Twins and Siblings Reared Together and Apart.—In an earlier issue of this *Journal*² I summarized the six or seven converging arguments which can be adduced in support of the inheritance of general ability. The most logical method of investigating such a problem is to keep each of the two variables constant in turn, and compare the results. Let us, therefore, take measurements first for children of identical heredity brought up in different environments and secondly for children of different heredity brought up in the same environment.

In the paper just cited, I gave correlations obtained originally from surveys in the London schools, and supplemented them by further data collected by Miss Conway, who had been responsible for the final computations. Thanks to numerous correspondents, she has since been able to increase the number of cases, particularly for the small but crucial groups of monozygotic twins reared together or apart. The total numbers now amount to 984 siblings, of whom 131 were reared apart; 172 dizygotic or two-egg twins, all reared together; 83 monozygotic or one-egg twins reared together, and 21 reared apart.³ By way of contrast, she has also secured data for 287 foster children.

¹ *Behaviourism* (1931), pp. 99f. Watson goes on to guarantee that "given my own world to bring them up in," he could train any healthy infant to follow any type of profession—"doctor, lawyer, artist, regardless of abilities or ancestors." Without going so far as this, Dr. Blackburn, Dr. Fleming, Dr. Heim, and a large number of sociological writers, appear to accept the general behaviourist view; but it should be noted that even Watson slipped in a few reservations which his more ardent disciples commonly omit.

So far as individual psychology is concerned—apart from the discredited claims of the Iowa school—no new facts have been responsible for this remarkable change of view: it seems rather to be an incidental symptom or consequence of an equally remarkable change in the general climate of opinion. In psychology as in politics, the pendulum of fashion swings to and fro; and the vacillations roughly synchronize. During the nineteenth century, the associationists preached an egalitarian doctrine, and three reform bills were passed. Then the close of the century witnessed a reaction, and we ourselves are witnessing the counter-reaction. An excessive emphasis on heredity has now been succeeded by an equally excessive emphasis on environment. Apparently it is difficult to give due weight simultaneously to each.

² "Ability and Income," *Brit. J. Educ. Psych.*, XIII, 1943, pp. 89-91.

³ Of the monozygotic twins, only nineteen were found in London; and, owing to the distances involved, we have been obliged to depend for measurements of the rest either on research-students or on local teachers and doctors (to whom we must extend our sincerest thanks). As a result, the correlations for this group may have been somewhat reduced. There is a natural prejudice against separating twins, especially if their sex is the same; and we should like to repeat our appeal for further cases. Although the handful of monozygotic twins reared apart is decidedly small (and it is the outcome of a quest that has lasted for over forty years), the differences between the correlations for this group and the rest are for the most part statistically significant.

The figures for head-length, head-breadth, and eye-colour are based on much smaller numbers in every batch. Eye-colour (assessed by the methods described in my paper in the *Eugenics Review*, XXXVII, 1946, pp. 149f.) was added because, of all readily observable traits, it is immune from environmental influence.

The correlations are set out in Table I. Since one or two writers apparently think that the figures obtained by American investigators imply different conclusions from those that I have drawn, I have also included the correlations obtained by Newman, Freeman, and Holzinger (15).¹

TABLE I
CORRELATIONS BETWEEN TESTS OF MENTAL, SCHOLASTIC AND PHYSICAL MEASUREMENTS.

Measurement	A—BURT AND CONWAY						B—NEWMAN, FREEMAN & HOLZINGER.		
	Identical Twins reared to- gether	Identical Twins reared apart	Non- identical Twins reared to- gether	Sib- lings reared to- gether	Sib- lings reared apart	Un- related child- ren reared to- gether	Identical Twins reared to- gether	Identical Twins reared apart	Non- identical Twins reared to- gether
MENTAL (INTELLIGENCE)									
Intelligence :									
Group Test	·944	·771	·542	·515	·441	·281	·922	·727	·621
Individual Test ..	·921	·843	·526	·491	·463	·252	·910	·670	·640
Final Assessment ..	·925	·876	·551	·538	·517	·269	—	—	—
SCHOLASTIC									
General Attainments	·898	·681	·831	·814	·526	·535	·955	·507	·883
Reading and Spelling	·944	·647	·915	·853	·490	·548	—	—	—
Arithmetic	·862	·723	·748	·769	·563	·476	—	—	—
PHYSICAL									
Height	·957	·951	·472	·503	·536	·069	·981	·969	·930
Weight	·932	·897	·586	·568	·427	·243	·973	·886	·900
Head Length	·963	·959	·495	·481	·536	·116	·910	·917	·691
Head Breadth	·978	·962	·541	·507	·472	·082	·908	·880	·654
Eye Colour	1·000	1·000	·516	·553	·504	·104	—	—	—

As regards intelligence, the outstanding feature of the table is the high correlation between the assessments for identical twins even when they have been reared apart : it is almost as high as the correlation between two successive testings for the same individuals. Between non-identical twins the resemblances (at any rate with our own data) are not much closer than those between ordinary brothers and sisters. Nevertheless, environment is not entirely without effect, particularly when the assessments have been obtained by written tests applied

¹ Dr. Heim, referring to the American inquiry, states that " when young monozygotic twins are separated . . . the differences between their scores are as great as those between unseparated dizygotic twins." But it will be seen that, in point of fact, both with the group test (Otis) and with the individual test (Stanford-Binet) the figures there given for the separated monozygotic twins are appreciably higher than those for the unseparated dizygotic twins, even though their figures for the latter are larger than those of most other investigators.

The figures obtained for twins in the most recent and extensive studies of twins carried out in Great Britain seem in the main to agree with our own. Herman and Hogben report with the Otis group test a correlation of 0·66 for twins of like sex and only 0·53 for twins of unlike sex : if we suppose that about half those of like sex were non-identical, this suggests a figure of about ·80 for the identical twins (12). Maxwell analysed data obtained with group tests for 468 twins during the Scottish Survey, and found correlations of 0·73 for twins of like sex and 0·63 for twins of unlike sex : as he observes, the latter value is " a little higher than that found in most other studies " (19).

to whole groups. The effect is obvious when we compare the correlations for children reared together and children reared apart. And it might be thought that in the correlations obtained from unrelated children reared in the same homes we have a direct indication of its actual amount. In all probability, however, such correlations mainly reflect the method of placement: a dull or defective orphan would not be boarded out with a highly intellectual family.

The figures for physical measurements, at least in our own data, show very similar trends: with the American data the correlations are somewhat higher, but the disparity is seldom large.¹

The results obtained for the scholastic tests, both in the American inquiry and in our own, present a striking contrast. In our own inquiry the correlations for siblings and non-identical twins reared together are actually higher than those for the identical twins who have been reared apart. And it may be instructive to note that the correlations which are most conspicuously increased by similarity of home environment are those for verbal or literary attainments; those for arithmetical attainments are, if anything, increased more by similarity of genetic constitution.

Figures like the foregoing provide ample evidence that individual differences in general intelligence are in part at least inherited, and that they are affected by environmental differences much less than are school attainments. However, the mere fact of hereditary influence the more sober critics do not deny. What they question is whether its amount is really large enough to be of any practical consequence either in the sphere of education or in later civic life.

Now I believe that a good deal of the difficulty arises because both the opponents of mental inheritance and its advocates still cling to wholly out-of-date notions of what is to be understood by such a phrase. Terms like heredity and variation, which played such crucial roles in the theories of Darwin, Spencer and the earlier biometricians, continue to be used by modern biologists, but their implications have radically changed. Moreover, the few educationists who appreciate the relevance of this change seem to be quite uncertain how far the newer theories have undermined the older inferences of the Galton-Pearson school.

The Hypothesis of Multifactor Inheritance.—Galton at the very outset of his work noted that in nearly all mental characteristics the observable differences between individuals are differences of degree rather than of kind, and proposed a scale of continuous variation in place of the traditional schemes of discontinuous types. Now, during the first two decades of the century, both the advocates of the new Mendelian hypothesis, and its opponents, originally supposed that the particulate theory of heredity, and the basic principle of segregation, were incompatible with continuous variation in an inheritable trait. Thus, Pearson and the biometric school contended that, even if true, the Mendelian hypothesis must be exceedingly limited in its application, and could have little or no bearing on normal psychology. On the other hand, the earlier Mendelians, De Vries, for example, believed that, since the Mendelian mechanism must underlie all forms of inheritance, no continuous variations could ever be inheritable; and this argument is still adduced by those who reject the inheritability of intelligence, because (so they assume) the very fact that variations in intelligence are continuous shows that they are produced by purely environmental agency.

¹ The high correlation for physical measurements obtained by Newman and his colleagues with non-identical twins is a little surprising. Lauterbach's figures agree more closely with my own. His correlations for twins of like and unlike sex are, for height, 0.80 and 0.53; for weight, 0.89 and 0.50 (*Genetics*, X, 1925, pp. 525-568).

The correlations are set out in Table 1. Since one or two writers apparently think that the figures obtained by American investigators imply different conclusions from those that I have drawn, I have also included the correlations obtained by Newman, Freeman, and Holzinger (15).¹

TABLE 1
CORRELATIONS BETWEEN TESTS OF MENTAL, SCHOLASTIC AND PHYSICAL MEASUREMENTS.

Measurement	A - BURT AND CONWAY						B- NEWMAN, FREEMAN & HOLZINGER.		
	Identical Twins reared to- gether	Identical Twins reared apart	Non- identical Twins reared to- gether	Sib- lings reared to- gether	Sib- lings reared apart	Un- related child- ren reared to- gether	Identical Twins reared to- gether	Identical Twins reared apart	Non- identical Twins reared to- gether
MENTAL (INTELLIGENCE)									
Intelligence :									
Group Test944	.771	.542	.515	.441	.281	.922	.727	.621
Individual Test921	.843	.526	.491	.463	.252	.910	.670	.640
Final Assessment . .	.925	.876	.551	.538	.517	.269			
SCHOLASTIC									
General Attainments	.898	.681	.831	.814	.526	.535	.955	.507	.883
Reading and Spelling	.944	.647	.915	.853	.490	.548			
Arithmetic862	.723	.748	.769	.563	.476			
PHYSICAL									
Height957	.951	.472	.503	.536	.069	.981	.969	.930
Weight932	.897	.586	.568	.427	.243	.973	.886	.900
Head Length963	.959	.495	.481	.536	.116	.910	.917	.691
Head Breadth978	.962	.541	.507	.472	.082	.908	.880	.654
Eye Colour	1.000	1.000	.516	.553	.504	.104			

As regards intelligence, the outstanding feature of the table is the high correlation between the assessments for identical twins even when they have been reared apart : it is almost as high as the correlation between two successive testings for the same individuals. Between non-identical twins the resemblances (at any rate with our own data) are not much closer than those between ordinary brothers and sisters. Nevertheless, environment is not entirely without effect, particularly when the assessments have been obtained by written tests applied

¹ Dr. Heim, referring to the American inquiry, states that " when young monozygotic twins are separated . . . the differences between their scores are as great as those between unseparated dizygotic twins." But it will be seen that, in point of fact, both with the group test (Otis) and with the individual test (Stanford-Binet) the figures there given for the separated monozygotic twins are appreciably higher than those for the unseparated dizygotic twins, even though their figures for the latter are larger than those of most other investigators.

The figures obtained for twins in the most recent and extensive studies of twins carried out in Great Britain seem in the main to agree with our own. Herman and Hogben report with the Otis group test a correlation of 0.66 for twins of like sex and only 0.53 for twins of unlike sex : if we suppose that about half those of like sex were non-identical, this suggests a figure of about .80 for the identical twins (12). Maxwell analysed data obtained with group tests for 468 twins during the Scottish Survey, and found correlations of 0.73 for twins of like sex and 0.63 for twins of unlike sex : as he observes, the latter value is " a little higher than that found in most other studies " (19).

to whole groups. The effect is obvious when we compare the correlations for children reared together and children reared apart. And it might be thought that in the correlations obtained from unrelated children reared in the same homes we have a direct indication of its actual amount. In all probability, however, such correlations mainly reflect the method of placement: a dull or defective orphan would not be boarded out with a highly intellectual family.

The figures for physical measurements, at least in our own data, show very similar trends: with the American data the correlations are somewhat higher, but the disparity is seldom large.¹

The results obtained for the scholastic tests, both in the American inquiry and in our own, present a striking contrast. In our own inquiry the correlations for siblings and non-identical twins reared together are actually higher than those for the identical twins who have been reared apart. And it may be instructive to note that the correlations which are most conspicuously increased by similarity of home environment are those for verbal or literary attainments; those for arithmetical attainments are, if anything, increased more by similarity of genetic constitution.

Figures like the foregoing provide ample evidence that individual differences in general intelligence are in part at least inherited, and that they are affected by environmental differences much less than are school attainments. However, the mere fact of hereditary influence the more sober critics do not deny. What they question is whether its amount is really large enough to be of any practical consequence either in the sphere of education or in later civic life.

Now I believe that a good deal of the difficulty arises because both the opponents of mental inheritance and its advocates still cling to wholly out-of-date notions of what is to be understood by such a phrase. Terms like heredity and variation, which played such crucial roles in the theories of Darwin, Spencer and the earlier biometricians, continue to be used by modern biologists, but their implications have radically changed. Moreover, the few educationists who appreciate the relevance of this change seem to be quite uncertain how far the newer theories have undermined the older inferences of the Galton-Pearson school.

The Hypothesis of Multifactor Inheritance.—Galton at the very outset of his work noted that in nearly all mental characteristics the observable differences between individuals are differences of degree rather than of kind, and proposed a scale of continuous variation in place of the traditional schemes of discontinuous types. Now, during the first two decades of the century, both the advocates of the new Mendelian hypothesis, and its opponents, originally supposed that the particulate theory of heredity, and the basic principle of segregation, were incompatible with continuous variation in an inheritable trait. Thus, Pearson and the biometric school contended that, even if true, the Mendelian hypothesis must be exceedingly limited in its application, and could have little or no bearing on normal psychology. On the other hand, the earlier Mendelians, De Vries, for example, believed that, since the Mendelian mechanism must underlie all forms of inheritance, no continuous variations could ever be inheritable; and this argument is still adduced by those who reject the inheritability of intelligence, because (so they assume) the very fact that variations in intelligence are continuous shows that they are produced by purely environmental agency.

¹ The high correlation for physical measurements obtained by Newman and his colleagues with non-identical twins is a little surprising. Lauterbach's figures agree more closely with my own. His correlations for twins of like and unlike sex are, for height, 0.80 and 0.53; for weight, 0.89 and 0.50 (*Genetics*, X, 1925, pp. 525-568).

Now, in spite of their undoubted importance for genetic and agricultural research, analogies drawn from the study of domesticated animals and plants may be highly misleading when we turn to human genetics. Very naturally, the characters that first caught the eye of the Mendelian experimentalist were qualitative traits, attributable each to some single factor or 'gene' which produces its own visible and distinctive effect. But there is no reason why genes should not exist whose separate manifestations evade our present methods of discrimination—systems of polygenes, and whose effects are *small, similar, and cumulative*.¹ If the number affecting the same trait were large, the result would be that observable variations in that trait would appear continuous, and the frequency-distribution of the measurements would approximate to the normal curve.² This is fully in keeping with the conclusions reached by Galton and his followers. It may be shown, says Galton, "that the distribution of human qualities and faculties (qualities like height and head-length, faculties like strength, visual acuity, or general ability) is approximately normal" (9, p. 32; (10, pp. 59, 201).

Manifestly it is impossible to check the existence of such genes by direct Mendelian methods; but, with the aid of statistics, we can discover whether the apparent effects are in accordance with Mendelian principles. Suppose, then, that a child's endowment of intelligence is dependent, not on a single pair of genes, but on many such pairs, each segregating in the usual fashion, and all affecting the same observable trait; and suppose too that one member of each pair (designated by a capital letter) would, if present, *add* a small quantity to the net result, while the other (designated by a small letter) would *deduct* an equal quantity. Then, for any given individual (or 'phenotype'), the total mount of intelligence would be proportional to the number of capital letters specifying the 'genotype.' Hence, if there were only three pairs of genes, the brightest individual would have a genetic constitution represented by *AABBCC*, the dullest a constitution represented by *aabbcc*, and the average person a constitution represented by *AaBbCc*. Assuming that mating is random and that there is no 'dominance,' the frequency of each genotype could be deduced by expanding the product $(A+a)^2 (B+b)^2 (C+c)^2$: it would, in fact, be proportional to the binomial coefficients, 1, 6, 15, 20, 15, 6, 1. With n such pairs of genes there would be $2n+1$ classes. And, as n increases, the binomial distribution will approach

¹ The possibility of multi-factor inheritance was mentioned by Mendel in his discussion of the colouring of white, red, and purple flowering beans. It was first demonstrated by H. Nilsson-Ehle in hybridization experiments on oats and wheat (*Kreuzungsuntersuchungen an Hafer und Weizen*, 1909); and the cardinal principles were elucidated more fully by E. M. East in studies of the corolla-length in the tobacco-plant ('Size-inheritance in *Nicotiana*' *Genetics*, I, 1915, pp. 164-176). The first to point out the importance of such a theory for human genetics appears to have been C. B. Davenport ('Inheritance of Stature,' *Genetics*, II, 1917, pp. 313f.). The number of genes which the theorist may legitimately postulate is now known to be far larger than was formerly thought: the banana-fly, *Drosophila*, is estimated to possess between 5,000 and 10,000; and man may have six times as many.

² For those who are not familiar with recent work in genetics, a brief explanation may be helpful. H. G. Wells, in one of his short stories, tells how an engaged couple hailing from North Wales—a Mr. Price-Jones and a Miss Evan-Roberts—plume themselves on bearing the family names of both their fathers and their mothers. But, they ask, how are they to christen their prospective children? The minister who is to marry them suggests that each child should take *one* surname from the male parent and *one* from the female, and that a coin should be tossed to decide the choice. Now let us apply the same principle to the case where a Mr. Price-Jones had married a Miss Price-Jones: the possible names for the children would be Price-Price, Price-Jones, Jones-Price, and Jones-Jones. This is exactly parallel to the way single genes are transmitted. Put *A* for Price and *a* for Jones. Then, when Mr. *Aa* marries Miss *Aa*, the possible recombinations are *AA*, *Aa*, *aA*, and *aa*: since *Aa* and *aA* are equivalent, the resulting proportions given by the toss will be 1:2:1.

more and more closely to the normal curve. But this, as we shall see, is only part of the story.

The Frequency-distribution for the General Population.—Modern critics of the Galtonian view usually start by attacking the theory of normal distribution. Dr. Heim, for example, assures us that it is a sheer assumption, "though not explicitly recognized as an assumption"; quite unwarrantably (she says) it has got "hailed as a scientific discovery, despite the fact that frequency distributions depend mainly on the system of scoring adopted." Mr. Richmond makes much the same point. To ensure this "a priori principle" (he says) the psychometrist "tinkers with the test material"; as a result "measurements are normally distributed, simply because the test has been so constructed that they must be so distributed" (11, 18).

Such arguments betray a singular indifference to the facts. In this country the first attempts to secure objective evidence about the distribution of test measurements were those made during my surveys of London schools. The chi-squared test was applied; and (as I pointed out at the time) the results disclosed quite plainly that such measurements are *not* distributed in exact conformity with the normal curve. The most conspicuous departure appeared in the lower tail of the curve, where, owing to an excess of dull and defective pupils (by no means invariably of a pathological type), the frequencies were much larger than the expected values.¹ When the defectives are omitted, then the resulting curve approximates more nearly to the normal, though the fit is still far from perfect. This *approximate* normality (which was all that Galton claimed) is thus not 'an a priori assumption' but an empirically demonstrated fact.

On examining the frequency curves for intelligence, therefore, we seem compelled to envisage two kinds of inheritance—unifactor inheritance and multifactor inheritance. If I may repeat what I have said elsewhere, "both the form of the distribution and the correlations obtained are very much what we should theoretically expect were these graded measurements, mainly though not wholly, determined by a very large number of similar genes; while in certain instances and in certain forms (as independent evidence from pedigrees suggests) mental deficiency may occasionally act like a dominant, or, still more frequently, like a recessive, and in some even be sex-linked": in this double mode of transmission, so I suggested, the inheritance of intelligence seems to resemble the inheritance of stature (3, p. 81). Moreover, as with stature so with intelligence, the observable measurements are in some degree modified by non-heritable influences. In the case of stature, the excessive frequency of very short persons is due partly to single genes (as with the achondroplastic dwarf, where the condition is dominant, and the ateleiotic dwarf, where it is apparently recessive), partly to environmental and pathological causes (as with rachitic or undernourished children), and sometimes to both (as with the cretin); and precisely the same types of causation are traceable in the dull and mentally deficient.

¹ A typical curve is that printed by Mayer Gross, Eliot Slater, and Martin Roth in their recent textbook on *Clinical Psychiatry* (1954, p. 56): the diagram is reproduced from one of my earlier surveys and based on over 3,000 cases: the irregularities are clearly visible.

Mr. Richmond cites as an example of 'tinkering' the revised version of the Binet-Simon Scale. But the tests were standardized with no reference whatever to normality: the assumption made was that, between the ages of 5 and 12, the annual increments are approximately equal. With properly constructed group tests, the items are selected (often by elaborate scaling techniques, such as paired comparison or its equivalents) so as to increase more or less uniformly in difficulty. Even in mechanical tests like erasing o's and e's in a page of pied print, where there can be no suspicion of 'tinkering with the scale,' the distributions are still approximately normal.

The Frequency-distributions for Parents and Siblings.—The possibility of polygenic determination was not overlooked by the biometric school. Galton himself was convinced that "inheritance may be described as largely, if not wholly 'particulate'" (10, p. 7). And Karl Pearson carried out a theoretical study of the statistical consequences of multifactorial inheritance (16). He concluded, however, that the correlations actually observed both between parents and their offspring and between children and their own brothers or sisters were far too high to be explicable by any such hypothesis. But, as now seems plain, his deductions were partly invalidated by certain untenable assumptions and several undue simplifications. To begin with, he tacitly assumed that dominance would be complete; furthermore, though keenly aware of the facts of assortative mating, he failed to make correct allowance for its influence; and above all, like most of the earlier biometricians, he failed to recognize the clear distinction between the causes of inheritable variation and their observable effects, between the carriers of heredity and the manifestations of heredity, in short, between what is conveniently called the 'genotype' (the hereditary determinants considered as a system typical of certain individuals) and the 'phenotype' (the kind of individual organism eventually produced by the interaction of the genotype with its particular environment); and it is a failure to recognize the same distinction that is largely responsible for the misconceptions and criticisms which the genetical psychologist encounters to-day.

The examination of the bivariate distributions is greatly simplified if we work with grouped frequencies. It is not difficult to show that, if a large number of genes combine, in the manner described above, to determine the measurements for two related members in a random sample of families (e.g., for parents and their children or for children and their sibs), and the measurements are suitably grouped to yield classes instead of continuous variates, then the frequencies to be expected will be similar to those deducible from a single pair of genes, for which the hybrid state (Aa or aA) is intermediate. Such frequencies, of course, can be readily computed by applying the ordinary principles of probability. The detailed values for multifactor inheritance have, in fact, been deduced by Fisher in his classical paper on 'The Correlation between Relatives on the Supposition of Mendelian Inheritance' (7): a non-technical account will be found in (8).

To ascertain how far the actual results for general intelligence conform with those which are required by the multifactor hypothesis, I have collected assessments for a 1,000 pairs of sibs, representing, so far as possible, a random selection of the London school population.¹ At the same time I have endeavoured, though with poorer success, to secure assessments for at least one parent. Since these proved obtainable for only 954 cases, the analysis has to be limited to this smaller number. On the basis of the measurements, the children were divided

¹ The inquiry was limited to children between the ages of 8 and 13, and was based primarily on verbal and non-verbal tests of intelligence. The actual measurements were transformed into standard scores (i.e., deviations divided by the standard deviation for each age); and these scores in turn were converted to terms of an I.Q. scale with a standard deviation of 15. Thus, the dividing lines for the three groups are approximately I.Q.'s of 90 and 110. Borderline cases were specially investigated in the light of the teachers' reports, and doubts resolved by individual testing. For the assessments of the parents we relied chiefly on personal interviews; but in doubtful and borderline cases an open or a camouflaged test was employed. The entire set of data on which the following tables are based were derived from four successive surveys carried out with the assistance of Miss Pelling, Mr. Seymour, Miss Richardson, and Miss Howard respectively. The methods adopted were slightly different in each; and the last was the most accurate. But, so far as the grouped frequencies are concerned, the results disclose no significant changes; hence, it seems legitimate to lump the whole series together for purposes of the present analysis.

into three groups—bright, average, and dull—in the proportions 1:2:1; and a similar classification was adopted for the parents. The percentages we should expect for the bivariate distribution, based on the triple assumption of random mating, Mendelian segregation, and no tendency to dominance, are shown below in Tables IIA and IIIA. They are, it will be noted, in the proportions 1, 1, 0; 1, 2, 1; 0, 1, 1 for parent and child, and 9, 6, 1; 6, 20, 6; 1, 6, 9 for pairs of sibs. On calculating the product-moment correlation for each hypothetical table, the value will be found to be exactly 0.500.

The observed frequencies, also reduced to percentages, are shown in Tables IIB and IIIB: (the perfect symmetry of the latter results from the procedure regularly followed in constructing a table for an intra-class correlation). It will be seen that the observed proportions agree tolerably well with the hypothetical; and, as we shall learn in a moment, the divergences themselves are very much what we should anticipate. The actual correlations, computed from the original data, were, for parent and child, 0.481, and for sibs 0.507 (computed from the pooled frequencies tabulated below, the values would be slightly different owing to the 'coarse grouping').

TABLE II
BIVARIATE DISTRIBUTIONS FOR PARENTS AND THEIR CHILDREN.

Parents	A.—THEORETICAL FREQUENCIES				B.—OBSERVED FREQUENCIES.			
	Children			Total	Children			Total
	Bright	Average	Dull		Bright	Average	Dull	
Bright	12.5	12.5	0.0	25.0	10.8	12.3	1.9	25.0
Average	12.5	25.0	12.5	50.0	13.4	26.5	10.1	50.0
Dull	0.0	12.5	12.5	25.0	0.8	11.2	13.0	25.0
Total	25.0	50.0	25.0	100.0	25.0	50.0	25.0	100.0

TABLE III
BIVARIATE DISTRIBUTIONS FOR SIBLINGS.

Children	A.—THEORETICAL FREQUENCIES.				B.—OBSERVED FREQUENCIES.			
	Children			Total	Children			Total
	Bright	Average	Dull		Bright	Average	Dull	
Bright	14.1	9.4	1.5	25.0	14.7	8.2	2.1	25.0
Average . . .	9.4	31.2	9.4	50.0	8.2	34.7	7.1	50.0
Dull	1.5	9.4	14.1	25.0	2.1	7.1	15.8	25.0
Total	25.0	50.0	25.0	100.0	25.0	50.0	25.0	100.0

A perfect agreement between the observed frequencies and the theoretical cannot possibly be expected, since there must be numerous unavoidable influences, tending partly to increase and partly to diminish the apparent correlation. (i) To begin with, like all mental measurements, assessments for intelligence, however scrupulously checked and adjusted, are in some degree distorted by the *unreliability* of the methods available. The best estimate for the reliability coefficient is 0.916. If we apply the usual correction for unreliability, the observed values would be raised to 0.525 and 0.554 respectively. (ii) But, as we have seen, the most punctilious attempts to assess 'innate ability'

(itself a purely hypothetical quantity) cannot entirely escape the effects of different *environmental conditions*; and, of course, for members of the same family the effects must generally tend in the same direction. How far this may have augmented the apparent correlation it would be hard to say: but the increase must almost certainly have been smaller than the decrease due to unreliability.¹

(iii) The most elusive tendencies to allow for are those of dominance and assortative mating. Were *dominance* complete, the expected correlations would be altered to $\frac{q}{1+q}$ and $\frac{1+3q}{4(1+q)}$ respectively, where p^2 , $2pq$, and q^2 denote the proportions of pure dominants, mixed dominants, and pure recessives respectively, and $q+p=1$. Thus, the effect of dominance is once again to lower the apparent correlations; but, unlike that of unreliability, it lowers them by widely differing amounts. Now the initial classification we have adopted makes $p=q=\frac{1}{2}$. Substituting this value in the fractions given above, we obtain, for the expected correlation between parents and their children, a coefficient of $1/3$, that is 0.333, and for the expected correlation between children and their brothers or sisters a coefficient of $5/12$, that is 0.416.² The observed values are significantly higher.

(iv) What then can have raised the absolute values to this high level? The most likely answer is *assortative mating*. How then can its presence be verified and its influence assessed? One of its calculable results would be to increase the variance of the younger generation. Now, if we may trust our rather crude measurements, the variance of the parents is only 12.3 I.Q., whereas (in virtue of the mode of standardization) the variance of the children is 15 I.Q. This is tantamount to an increase in the filial generation of about 22 per cent. Spouses, it appears, prefer partners whose intelligence in some degree resembles their own. The actual amount of the resemblance can be estimated by calculating the correlation between husband and wife. In the earlier surveys it was well over 0.40; in the later somewhat below.

Now, as we have seen, dominance, like unreliability, tends to reduce the correlations between parent and child and between one sib and another; but, unlike unreliability, it reduces them to different extents. On the other hand, the effect of assortative mating, like that of similar environment, is to increase the correlations, and to increase them by amounts that make them more nearly equal. The net result can be estimated, if we treat the contributory variances as additive components, and then apply the ordinary principles of factor analysis. A little calculation indicates that the ultimate effect of assortative mating would be to *add a small amount to both correlations*, viz., in the present

¹ The critic usually supposes that intelligence-tests are considerably affected by cultural differences in the testees' environment. But, if the tests have been properly constructed and their pronouncements properly checked and adjusted, such effects are almost negligible. The influence of unhygienic conditions in early infancy is a more likely source of error, for which it is difficult to allow. The statement in the text was based on indirect attempts to estimate the upper limit for environmental influences by methods which need not be detailed here.

² These are the theoretical values deduced by Pearson in the paper already cited (16). He rejected them, and with them the assumptions on which they were based, because they fall far below the correlations he had empirically obtained for numerous traits showing continuous variation. Yule, however, pointed out that if, instead of postulating complete dominance (as Pearson had tacitly done), we postulate complete absence of dominance, both the theoretical values would be raised to 0.500, and this would accord far better with Pearson's own figures (*Report of Conference on Genetics*, 1906). Yule's assumptions make the relevant conditions even simpler than Pearson's: but in view of recent results, it seems pretty certain that they are far more complex.

instance about 0.15 and 0.09. With complete dominance, this would raise the theoretical values from 0.333 and 0.416 to about 0.48 and 0.52 respectively. And these figures tally reasonably well with those observed.¹

The Relative Importance of Heredity and Environment.—We now reach our final problem: what proportion of the total variance shown by the children is attributable to genetic conditions as contrasted with environmental? In recent discussions on this point, two important considerations are frequently ignored.

(i) If the observed correlation between parent and child is 0.481, we might infer that each parent contributes $0.481^2 = 23$ per cent. to the total variance. And if the mating were random, the two parents together would contribute $2 \times 23 = 46$ per cent. But since, as we have seen, there is a correlation of at least 0.40 between fathers and mothers, part of the influence of one parent must overlap with that of the other, and consequently should not be included twice. Making due allowance for the overlap, we may estimate the contribution deducible from the assessments for the two parents as about 45 per cent. at most. Now it is often inferred that the remainder of the variance must, therefore, be ascribed to non-inheritable factors, that is, to the influence of the environment. But with the mode of transmission we have assumed, not only the parents but also the grandparents and remoter ancestry must contribute something to the variance. A simple algebraic deduction from the postulates of multifactorial inheritance will show that the total effect of parentage and ancestry may be directly measured by the correlation between sibs. The observed correlation, it will be remembered, was 0.507. According to our findings, therefore, about 51 per cent. of the variance must be contributed by such factors.

(ii) But, even so, it would be quite mistaken to assign the whole of the residue (49 per cent.) to environmental influences. By an odd paradox, not only the similarity between siblings, but also their differences are largely the outcome of their genetic constitution. Thus, arguing from Mendelian principles, we should definitely anticipate a frequent lack of resemblance between one sib and another owing to the segregation of those factors in respect of which the parents are heterozygous. After computing a rough estimate for this additional contribution, I calculate that in all at least 75 per cent. of the entire variance must be due to genetic influences, probably far more.²

It must be frankly owned that, with a sample covering under a thousand cases, the somewhat speculative balancing of the accessory factors that affect the correlations here obtained can make no pretence to be either accurate or

¹ It might be suggested that the resemblance between brothers and sisters appears greater than that between parents and their children, because children of the same family are brought up together and may even go to the same school. Similarity of schooling might no doubt affect the correlations for cultural and educational tests, as indeed the figures in Table I suggest; but (for the reasons given above) it cannot appreciably affect the results obtained for intelligence. Certainly the assessments for siblings who have gone to the same schools reveal no higher correlation than the assessments for those who have gone to different schools.

² The theoretical considerations on which such calculations should be based are clearly set out by Fisher in the paper already cited (7). It may be observed that the figure for the residual contribution which we have thus reached, namely 25 per cent., would imply a correlation with the conditions causing it amounting to $\sqrt{0.25} = 0.50$. But a direct calculation of the partial correlation between favourable environmental conditions and the calculation of the partial correlation between favourable environmental conditions and the assessments for intelligence proves to be well under this figure. Hence, the final figure reached above for genetic influence leans definitely to the conservative side. Fisher's formulae would subdivide the contributions to the total variance into (i) genotypes 49 per cent., (ii) dominance 28 per cent., (iii) assortative mating 19 per cent., leaving for (iv) environment only 4 per cent.

conclusive. My aim has rather been to adumbrate a line of reasoning that merits closer consideration and further research. But, even as it stands, the analysis I have made, supplemented by the other evidence that I have mentioned, seems to me to afford a strong corroboration for the view I have indicated, namely, that *human intelligence, like human stature, is determined largely though not wholly by multifactorial inheritance.*¹

Conclusion.—I have now reviewed the wide variety of evidence—observational, introspective, and experimental, biological, physiological, and statistical, bearing on our initial question. The results are mutually supporting; and, apart from certain minor modifications or extensions, seem abundantly to confirm the threefold hypothesis that I tentatively put forward over forty years ago in the forerunner of this *Journal*²: namely, that there is a general factor making for efficiency in all mental activities, that this factor is essentially cognitive or directive, and that the greater part of the individual variance found in this factor is attributable to differences in genetic constitution. This triple conclusion suggested a modernized formula for the abstract conception to which so many different writers had been led, viz., 'innate, general, cognitive ability.' If, therefore, we are to retain the word 'intelligence' as a technical term in psychology, this still seems the best definition.

III.—SUMMARY.

The main steps in the argument may be epitomized as follows:

- 1.—Evidence from different branches of psychology leads to the notion of a mental capacity that is (i) cognitive, (ii) general, (iii) innate.
- 2.—Each of these three characteristics has been amply verified by statistical research.
- 3.—As the history of the word shows, intelligence was a technical term put forward to designate a technical concept: and the meaning given it, implicitly or explicitly, by leading authorities from Cicero and the scholastics to Spencer, Galton, and Binet, suggests that it furnishes the most convenient name for the concept thus reached.

¹ As I suggested in my earlier paper, it is urgently desirable that similar methods should be employed to investigate the presence of a general cognitive factor in lowlier animals, and, if possible, to determine its mode of inheritance. The few researches so far carried out point to conclusions similar to those reached above. R. L. Thorndike has calculated correlations between tests of learning, strength of drive, etc., in albino rats, and finds a general factor of learning ability and two supplementary factors (*Genet. Psych. Monogr.*, XVII, 1935, pp. 1-70). Vaughn (*Comp. Psych. Monogr.*, XIV, 1937, pp. 1-41) and Tolman and others (*ibid.*, XVII, 1941, pp. 1-20) have also published tables of correlations for the performances of rats; their figures are fully consistent with the theory of a general factor, though the investigators themselves prefer an analysis in terms of overlapping group factors.

To test the hypothesis of multifactorial inheritance, Tryon has carried out experiments on maze-learning, which he regards as a test of general ability, and has repeated them with successive generations. He attempted first to secure two strains, of bright and dull rats respectively, by selective inbreeding. After seven selections and seven generations, he found practically no overlapping between the distribution-curves for bright and for dull. He then, as it were, reversed the procedure, crossing the two strains, and testing two further generations. It is true that the variance exhibited by the F_1 -generation seemed too great to be explained wholly by non-genetic influences, and much greater than would be expected had the method of inbreeding been successful. Yet, on the whole, as he contends, the results seem to support the multifactorial hypothesis (R. C. Tryon, "Genetic Differences in Maze-Learning in Rats," *Thirty-ninth Yearbook of Nat. Soc. Study of Education*, 1940, pp. 113f.; also E. G. Brody, "Genetic Basis of Spontaneous Activity in the Albino Rat," *Com. Psych. Monogr.*, XVII, 1942, No. 5.).

² See *J. Exp. Pedag.*, I, 1911, pp. 93f; Cf. also ¹.

4.—Apart from comparatively rare and abnormal variations, differences in intelligence as thus defined seem to depend on the combined action of numerous genes whose influence is similar, small, and cumulative—a hypothesis that is fully borne out by the frequency-distributions obtained for parents, siblings, and the population as a whole. And on this hypothesis not only the similarities between relatives but also their dissimilarities will be largely due to genetic factors.

5.—It is essential to distinguish between intelligence as an abstract component of the individual's genetic constitution (γ) and intelligence as an observable and empirically measurable trait (g). The evidence indicates that at least 75 per cent. of the measurable variance (based on carefully checked assessments) is attributable to differences in genetic constitution, and less than 25 per cent. to environmental conditions.

IV.—REFERENCES.

- (1) BURT, C. (1912): "The Inheritance of Mental Characteristics," *Eugenics Review*, IV, pp. 1-33.
- (2) BURT, C. (1924): "History of the Development of Psychological Tests," Ap. Board of Education, *Report on Psychological Tests of Educable Capacity*. (London: H.M. Stationery Office.)
- (3) BURT, C. (1935): *The Subnormal Mind*. (Oxford: The Oxford University Press.)
- (4) BURT, C. (1946): *Intelligence and Fertility*. (London: Hamish Hamilton.)
- (5) BURT, C. (1955): "The Meaning and Assessment of Intelligence" (Galton Lecture for 1955), *Eugenics Review*, XLVII, pp. 81-91.
- (6) CARMICHAEL, L. (ed). (1946): *Manual of Child Psychology*. (New York: Wiley.)
- (7) FISHER, R. A. (1918): "The Correlation between Relations on the Supposition of Mendelian Inheritance," *Trans. Roy. Soc. Edin.*, LII, pp. 399-434.
- (8) FISHER, R. A. (1919): "The Causes of Human Variability," *Eugenics Review*, X, pp. 213-220.
- (9) GALTON, F. (1869): *Hereditary Genius*. (London: Macmillan.)
- (10) GALTON, F. (1889): *Natural Inheritance*. (London: Macmillan.)
- (11) HEIM, A. W. (1954): *The Appraisal of Intelligence*. (London: Methuen.)
- (12) HOGBen, L. and HERMAN, L. (1933): "The Intellectual Resemblance of Twins," *Proc. Roy. Soc. Edin.*, LIII, pp. 105-129.
- (13) KIRMAN, B. H. (1952): *This Matter of Mind*. (London: Watts.)
- (14) NEWMAN, H. H., FREEMAN, F. N., and HOLZINGER, K. J. (1937): *Twins: A Study of Heredity and Environment*. (Chicago: University of Chicago Press.)
- (15) NEWMAN, H. H. (1942): *Twins and Super-Twins*. (London: Hutchinson.)
- (16) PEARSON, K. (1903): "On a Generalized Theory of Alternative Inheritance with Special Reference to Mendel's Laws," *Phil. Trans.*, CCIII, 53-87.
- (17) PIAGET, J. (1950): *The Psychology of Intelligence*. (London: Routledge and Kegan Paul.)
- (18) RICHMOND, W. K. (1953): "Educational Measurement: Its Scope and Limitations: A Critique," *Brit. J. Psych.*, XLIV, pp. 221-231.
- (19) SCOTTISH COUNCIL FOR RESEARCH IN EDUCATION (1953): *Social Implications of the 1947 Scottish Mental Survey*. (London: University of London Press.)
- (20) SPEARMAN, C. (1927): *The Abilities of Man*. (London: Macmillan.)
- (21) THOMSON, G. H. (1948). *The Factorial Analysis of Human Ability*. (London: University of London Press.)
- (22) WISSLER, C. (1901): "The Correlation of Mental and Physical Tests," *Psych. Rev.* Monogr. Supp., III, vi.

A STUDY OF SPECIFIC ABILITY AND ATTAINMENT IN MATHEMATICS

By DORIS M. LEE

(*University of London Institute of Education.*)

I.—*Need for present work.* II.—*Earlier investigations.* III.—*Present investigation.* IV.—*Analysis of test scores.* V.—*Factor analyses.* VI.—*Interpretation of results.* VII.—*Summary.* VIII.—*Appendix of tests of mathematical ability.* IX.—*Acknowledgments.* X.—*References.*

I.—NEED FOR PRESENT WORK.

THIS research was begun by the writer during a term of service in a large co-educational grammar school. Continuous observation of the differing academic achievements of individual children suggested a need to distinguish between ability (potential success) and attainment (actual success) in the study of any subject. The general results of an intelligence test seemed insufficient for this purpose, since a child of high intelligence did not always maintain a high level of attainment in all subjects. Beyond the level of estimating general intelligence, there seemed a need for some objective means of gauging the relevant specific ability of those whose attainment marks in a given school subject were continually low. As a basis for further exploration, *ability to succeed in a subject* was defined as the power to grasp and manipulate the fundamental notions required for its study. *Attainment in the study of a subject* was taken as the acquired degree of mastery over the complexities of its conventional structure, starting from the simplest beginnings and constantly proceeding to more difficult work. Under these definitions, ability is related to school work and is certainly necessary for achievement in it, for some grasp of fundamental notions (though not explicitly demanded) is implicitly contained in the study of any subject. Attainment in a subject entails a sequence of work which rests on a prior acceptance of certain basic elements, concepts and modes of thought. At the school level, it seemed probable that an assessment of the child's success in grasping these basic notions might prove a useful pointer to his likely success in the study of the corresponding subject and hence to an estimate of his ability in the subject itself. Since the writer was principally concerned with the teaching of Mathematics, it was decided to limit a first investigation to the quest for such a measure of ability in the field of school mathematical work.

II.—EARLIER INVESTIGATIONS.

Examination of available psychological literature revealed no study comparable to that envisaged above. Separate literature existed concerning mathematical ability and attainment, some of which suggested lines of approach for the investigation under consideration. Early tests of mathematical ability¹ had, in reality, judged attainment, but the later work of Hamley² provided a basis for mathematical tests which were unrelated to school work. A succession of small pieces of work³ which had followed suggested that such tests could function as mathematical ability tests. Hamley discussed the hypothesis that three stages could be distinguished in the process of mathematical reasoning : (1) Dividing the given material into classes, each determined by some unifying characteristic.

- (2) Detecting a prevailing order within each class.
- (3) Seeking out correspondences of relationships between the members of two or more classes.

He felt that the ability to succeed in the study of Mathematics could correspond to the ability to carry out these three processes of classification, order, and recognition of correspondences in dealing with the simple material which is basic to Mathematics, namely, arithmetical numbers, algebraic symbols and spatial figures. Hamley lists eight simple and six more complex tests⁴ designed to discover whether children could manipulate these basic processes of mathematical reasoning in terms of simple mathematical material, though the more complex tests contain much verbal material, and are to some extent allied to work which is taught in the secondary school mathematics courses. The tests were administered to a number of average elementary school children in the London area (average age 11.2 years; average I.Q. 102). The comments printed with the tests⁵ show no undue difficulty or unfamiliarity on the part of these testees.

Jenkins⁶ introduced an additional stage into this scheme of mathematical reasoning, when he suggested that the ability to recognise some constant variable within a group preceded the identification of that group as a class, in the sense defined by Hamley. He showed that, with this addition, Hamley's analysis⁷ of the dynamic processes of mathematical reasoning corresponded to Russell's delineation⁸ of the fundamental concepts of Mathematics in the more static field of Mathematical Philosophy. Jenkins constructed a battery of seventeen⁹ tests based on four processes of seeking out a variable, classification, order and correspondence in terms of mathematical material, in the hope of establishing a battery of tests for mathematical ability. Amongst these were included some of the simple tests originally devised by Hamley and some more complex forms weighted heavily with verbal material. In separate factor analyses of scores obtained from administering these tests to three-hundred and twenty-three grammar school boys and girls aged from 13 to 15 years, only isolated factors reflecting the processes of classification (in the case of both the boys' and girls' scores) seeking out an order (in the case of the boys' scores only) and of discovering correspondences (in the case of the girls' scores) were obtained. The analysis patterns seemed heavily obscured by a verbal factor (probably due to the excessive amount of verbal material in some of the tests) and by certain structural factors of number and space (probably occasioned by the lack of balance in the battery between processes of reasoning, on the one hand, and recognition of variable were found to be unsuitable and had to be discarded. Nevertheless, there was sufficient evidence to suggest that suitable tests of mathematical ability could be obtained by working along these lines.

The work of Hamley and Jenkins rests on the assumption that the fundamental notions of Mathematics are of two kinds:

- (1) The notions of arithmetical numbers, algebraic symbols and spatial figures. The study of Mathematics cannot proceed at all without these, for they provide the material with which the mathematician works. They also designate the subject field called Mathematics.

- (2) The notions of variable, class, order and correspondence. A *variable* is a quantity which can take different values, although its basic structure remains the same throughout; an even number is a *variable*, since there are many such numbers all having the property of exact divisibility by two. A *class* is a group of quantities having a common characteristic; a group of even numbers is thus a class, the common characteristic being the even-ness of the numbers,

or the property that each is exactly divisible by two. This common characteristic is also the 'variable' of that 'class.' Hence the notions of variable and class are very much interwoven. An *order* is obtained when quantities are arranged in a sequence according to a fixed rule or law; thus, even numbers can be arranged in ascending or descending order. It is worth noting that a given class has only one unifying characteristic or variable, but usually it is possible to arrange the class in several different orders. Thus, the notions of variable and class are more closely linked than those of variable and order, or than those of class and order. A *correspondence* is obtained when two ordered classes of quantities are placed side by side in such a way that each pair is bound by a law which can be enunciated; thus, the sequences of odd and even numbers show a correspondence in which each member of one class differs by one from the corresponding member of the other class. A correspondence is hence closely dependent on the orders of the two classes in much the same way as the notion of class depends on that of variable. The notion of correspondence is also linked to the notions of class and variable, since these are contained in the notion of order. The notions of variable and class could thus be considered basic or primary to the notions of order and correspondence. These notions of variable, class, order and correspondence play a large part in mathematical work, since most problems are solved by recognising classes amongst given data, creating orders within these classes, and picking out correspondences to lead to a unique conclusion. On the other hand, these notions are not confined to Mathematics alone. They may well provide a link between Mathematics and other subjects, though in other fields the conclusion is not always unique. In this light, a subject may be designated by the kind of material with which it deals, and not by the general manner of arranging that material to form the conventional subject pattern.

Philosophers¹⁰ have suggested other fundamental notions of Mathematics such as similarity or equivalence, function, inference, functional deduction and symbolism. A critical examination of these suggests that most of them are types of correspondence usually applicable in limited sections of the subject. For example, similarity is customarily associated with geometry, and function with algebra. The idea of correspondence is more general, and covers these more limited notions, as well as many others.

The notions of variable, class, order and correspondence can be seen in every phase of mathematical work, and hence can be regarded as fundamental notions required for the study of Mathematics. Together with the notions of arithmetical numbers, algebraic symbols and spatial figures, these provide the basis for tests of mathematical ability according to the definition of ability given in Section I. From the psychological angle, this is equivalent to the hypothesis that four stages can be distinguished in the dynamic process of mathematical reasoning, namely those of recognising a variable, classifying, ordering and recognising correspondences in sets of arithmetical numbers, algebraic symbols, and spatial figures.

III.- PRESENT INVESTIGATION.

The present investigation seeks to devise and evaluate separate batteries of tests for measuring ability and attainment in Mathematics. The evaluation of the test scores is expected to give some idea of how far the individual's ability for Mathematics is reflected in his attainment in school mathematical work, and to present some confirmation of the subjective structure of mathematical ability as laid out in Section II. Since Mathematics is mainly taught in

grammar schools, fullest measures of attainment are likely to be obtained by confining the investigation to this type of school only, though true tests of mathematical ability could be administered to any group of children, irrespective of differences in external factors which may obtain within the group. Again, from the limited point of view of a controlled investigation, working with grammar school children introduces a narrower and more homogeneous range of intelligence, at the same time permitting a more heterogeneous range of mathematical ability. While this allows examination of certain relations between the tests of mathematical ability and the intelligence tests, it may prevent a general intelligence factor from obscuring the resulting factor patterns. The presence of intelligence is acknowledged, but the emphasis of the investigation is towards constructing tests which may throw light on the nature of the specific abilities involved in the study of Mathematics at the school level. Some group factors may enter into the results, though the scope of the proposed ability tests is such as to minimise even these. The relation of some well-known group factors to the form of tests to be employed in the present investigation could well form the subject of a separate research.

The customary grammar school criterion of satisfactory attainment in Mathematics is a 'pass' in the subject at Ordinary Level in the General Certificate of Education Examination. The syllabus for this examination provides the backbone of the mathematical work for the five-year grammar school course, and hence seems a suitable framework upon which to build attainment tests for the present purpose. Scrutiny of this syllabus shows the three types of mathematical work falling under the headings of Arithmetic, Algebra and Geometry, respectively, and these three divisions correspond to the three kinds of material with which all Mathematics performs its functions. Within these three syllabus divisions, a total of fifty-three distinct topics are taught, including such items as Percentages in Arithmetic, Quadratic Equations in Algebra, and the group of Circle Theorems in Geometry. In order to obtain a measure of attainment in grammar school Mathematics, a battery of fifty-three attainment tests was accordingly constructed. Each attainment test was validated by a consensus of opinion amongst the Mathematics staff, in the manner described by Long and Sandiford,¹¹ and the following principles were offered for guidance:

- (1) Each test is one of principle rather than manipulation, and hence the foundation figures employed are to be simple in nature.
- (2) As far as possible, the items in each test shall be graded in order of difficulty.
- (3) In order to exhibit and sample the mathematical principle in the widest possible manner, the verbal element is to be kept to a minimum. Questions will also tend to be of the short type rather than the long.
- (4) Consideration of the number of tests indicates that brevity is desirable, but as many aspects of the principle as possible shall be included; no form will work all the tests at one sitting, but the encroachment of testing time on school mathematics periods must also be considered.

When agreement was reached according to these requirements, the time to be allowed for working the tests was discussed and decided in the light of the total experience of the group of staff members.

Since mathematical ability is to be judged independently of attainment in the subject, it seemed desirable to construct one set of tests of mathematical ability for use at all levels of the five-year grammar school course. This pre-supposes that the individual's ability for Mathematics remains the same as he progresses through the school, though increasing familiarity with the sort of material in which these tests are couched may tend to improve the scores of the

older children. Such a plan would also facilitate comparisons between the results obtained from factor analyses at the different age-levels. Following the sequence of work described in Section II, a battery of tests of mathematical ability was built by combining the four fundamental processes of recognition of variable, class, order and correspondences with the three basic types of material employed in the study of elementary Mathematics, these being arithmetical numbers, algebraic symbols and spatial figures. Separate tests were designed to estimate capacity for handling each of the four processes in terms of each of the three types of material making twelve tests in all. The combinations of processes and material are shown in Table I,* together with the code letters used to designate each test.

TABLE I

Three types of Material	FOUR BASIC PROCESSES			
	Recognition of Variable	Classification	Recognition of Order	Recognition of Correspondence
Arithmetical Numbers	V.N.	Cl.N.	O.N.	C.N.
Algebraic Symbols	V.Z.	Cl.Z.	O.Z.	C.Z.
Spatial Figures.	V.S.	Cl.S.	O.S.	C.S.

The tests† were based on simple evaluation of the four processes, and the form in which each process was tested remained the same for each type of material. All were developments of the simple tests originally outlined by Hamley in this field, though necessary gaps were filled in order to maintain a strict balance between the processes and the types of material in which the questions were couched. Table I shows that no single test exerted a greater influence than the remaining ones on the force of the battery as a whole, and the later emergence of casual structural factors of number and space may be minimised in this way. This use of straightforward tests also minimised the verbal material, and hence may lessen the presence of a heavy verbal factor in the resulting factor patterns. Preliminary versions of these tests were administered to three-hundred children at all levels of the five-year grammar school course. Items for the final test versions were validated by the method of Upper and Lower Thirds, and times for working the final tests were obtained. In order to check and identify the presence of general mental ability factors in any factor patterns which might result from this investigation, it was decided to include verbal and non-verbal intelligence tests. Consultation with the surviving co-author of the Fleming-Jenkins test "Mental Ability III"¹² suggested that equal portions of verbal and non-verbal material from this test could be adequately used over a five-year range, provided that shortened time limits were assigned for working the separate sections. A trial run with a group of sixth-form pupils showed that three-fifths of the original time-limits would be suitable, and these reduced estimates were recorded for use in the later administration of the tests. For experimental purposes, the whole Fleming-Jenkins test was included with the twelve mathematical tests, and the scores on the selected non-verbal and verbal portions were abstracted afterwards. This complete battery of tests

* Cf. HAMZA, M. (3).

† Samples of the tests are given in the Appendix of Tests.

permitted a division of the tests for two testing periods of forty-six and forty-five minutes, respectively, the first group being those couched in arithmetical and algebraic material and the second containing the spatial and general mental ability tests.

Consideration of possible experimental samples led to the use of all the pupils in the five-year School Certificate Course at the grammar school where the writer was a member of the Mathematics Staff, and this numbered more than one-hundred pupils at each year-level. Since the school was co-educational, both boys and girls could be included in the investigation. They worked together in the Mathematics classes and were taught by the same teachers. Thus, the examination of likely sex differences seemed a question for examination rather than deletion. Further, amongst the Mathematics staff, there was considerable unanimity on questions of teaching method and general educational outlook, and therefore the effect of a number of external factors on the attainment tests was also minimised. As regards the actual Mathematics teaching, syllabuses for forms in the same year-level were the same, and opportunity was given for every pupil to cover the ground for the year, the only modification for the less able child being differences in the initial methods of presenting the work. Thus, the numbers and conditions in the school seemed to provide a suitable setting for work of this nature. It was accordingly decided to administer the battery of ability tests at all levels of the five-year course in order to obtain an idea of the success of the tests at each stage. To avoid likely effects of discussion between testees, working at different times of day, working the tests in differing orders and other similar factors, it seemed reasonable to investigate the possibility of administering these tests to all pupils on the same day, at the same time and in the same order. The battery of ability tests was, therefore, administered to the whole five-year population in the school periods just before and after the mid-day break on a day following the close of the Summer Examination for the General Certificate of Education. A corresponding consideration of the battery of attainment tests led to a division into five groups, so that each test was administered to those pupils who had studied the corresponding mathematical work during that school year, the administration of all tests being carried out as the year came to its close. This to some extent minimised the likely differences of memory and time-recall, for each pupil had had the opportunity to learn the tested work within a limited time. A further arrangement enabled the attainment tests to be used as introductory tests for revision lessons, so that no work on the topic had taken place immediately prior to the administration of the test. Ready co-operation on all these points was obtained from the Head Master and staff of the school, all of whom very kindly helped with the final administration of the tests.

IV.—ANALYSIS OF TEST SCORES.

The scores for the two batteries of tests were recorded side by side for each child at each year-level. From the preliminary lists of scores, it seemed clear that both sets of tests had shown adequate discriminatory power within the limits of the sample employed. Further analysis of the score distributions and their inter-relations required investigation of two sources of disturbance which had had to be retained for the proper conduct of the investigation as a whole. At each year-level, the distributions included scores obtained from both boys and girls, and these distributions could not be used as single units without an investigation of the relevant differences between the boys' and girls' performances. Again, consideration of the attainment tests in relation to the school setting had demanded a grouping of the testees in year-levels rather than in

age-groups, and the variation in age in the separate year-levels ranged between twenty-five and twenty-nine months. It was possible that such ranges might influence correlations and hence interfere with the resulting factor patterns unless due care were taken to ensure that the likely influence of age was examined in the course of the analysis of the score distributions. Accordingly, the distribution of the boys', girls', and combined scores were examined separately for each test at each year-level. The number of significant differences between the boys' and girls' performances and the number of non-normal total distributions at each year-level were not found sufficient to interfere with the general analysis of the investigation. The boys' and girls' scores were, therefore, combined to make one single distribution for each test, and these distributions were used as a basis for the subsequent correlation and factor analysis work.

The average ages of the pupils in the separate year-levels increased by exactly one year at each successive stage; and, though this was a coincidence, it provided a satisfactory span of age-levels throughout the sample. However, since the test scores were grouped in year-levels, and not in age-groups, the likely influence of age could not be examined by any method of significant differences between the year-level groups. Accordingly, age was included in the variables for the correlation matrices at each year-level, in order that undue influence of age might be detected in the form of significant correlation coefficients between age and other variables. The square correlation matrices between age and the fourteen ability test variables were calculated for each year-level. The corresponding rectangular correlation matrices between the age and ability test variables and the attainment test variables were also obtained. The influence of age was strongest amongst the attainment tests at the fourth and fifth year-levels. These significant correlations with age were partialled out from the remaining correlations, so that the influence of age on the attainment test scores was considerably lessened. Apart from this one trend, the general pattern of the correlation matrices was of the form usually associated with work of this type.

Further expression of these correlations in the form of prediction equations was discarded in favour of factor analysis techniques, in the hope that the latter would reveal something of the structure of reasoning called into play by the ability tests, as well as the extent to which the same factors seemed to operate in the working of both kinds of tests. Since, in Section III, it was assumed that the ability to do Mathematics remained fairly constant throughout a pupil's school life, the correlation between an ability and an attainment test could justifiably be taken as an expression of the degree to which the factors underlying success on the ability test seemed to be called into play in the working of the attainment tests, and not just of the overlap between the two tests. Following from this, factor analyses at each year-level might throw light on the general relationship between the ability and attainment tests.

V.—FACTOR ANALYSES.

Factor analyses of the square correlation matrices from the fourteen ability tests were made for each year-level, in order to see whether the factors involved in the children's reasoning reflected the hypothetical structure of mathematical thinking outlined in Section II. Using Thurstone's Centroid Method, three factors were extracted at each year-level. Final factor-loadings were obtained by a system of two-by-two rotations based on the sign-pattern of the original centroid factors. In the case of the first, second and fourth year-levels, the third factor was small compared to the other two. In the case of the third and fifth year-levels, the second and third factors were both small.

Using the joint rectangular matrices of correlations between the mathematical ability and attainment tests at each year-level, an extension of the centroid method¹³ was employed in order to estimate how far the factors underlying performance on the ability tests also entered the working of the attainment tests. All factors were found to enter to some extent, and the degree to which they accounted for the correlations in the corresponding joint rectangular matrix varied from .10 to .80, though in general, the limits were .20 and .60, only isolated cases lying outside this smaller range. Significant loadings for all the factors entering the ability tests were found in twenty-nine out of fifty-three attainment tests. A further fifteen tests showed significant loadings on one or both major factors. The remaining tests also showed traces of all the factors.

VI.—INTERPRETATION OF RESULTS.

A general survey of the distributions arising from the administration of the tests of mathematical ability showed a movement from the lower half of the score range at the first-year level towards the upper half at the fifth-year level. Some evidence of skewness at the first and fifth-year levels suggested that the tests of mathematical ability were most suited to the second, third and fourth forms. Amongst the individual tests, those requiring recognition of a variable were found difficult at each year-level. This may be due to the fact that the recognition of variables is implicit rather than explicit in most mathematical work. Nevertheless, the general lay-out of the distributions resulting from the ability tests suggested that these had been suited to the sample and to the investigation as a whole. A similar consideration of the distributions arising from the attainment tests suggested that these had also been suited to the sample and to the whole investigation.

Later correlation of the scores on the ability tests with the Mathematics marks obtained by the testees in the examination for the General Certificate of Education gave coefficients varying between .13 and .57, though the tests couched in arithmetical numbers and algebraic symbols gave coefficients of the magnitude of .40 and .50.

Correlation of the scores from the tests of mathematical attainment with the later Mathematics marks obtained by the testees in the General Certificate Examination gave higher coefficients varying between .45 and .74. These positive correlation coefficients to some extent confirm the suitability of the two batteries of tests used in the present investigation. They suggest that the kind of mathematical ability defined and tested contributes to success in the conventional school mathematical work. The correlation coefficients between the attainment tests and the School Certificate marks are, in fact, similar to those obtained by Emmett and Wilmot,¹⁴ and by Peel and Rutter,¹⁵ in recent investigations into the prediction of School Certificate Performance. The tests of mathematical ability presented material according to Hamley's delineation of the stages of mathematical reasoning, and the resulting factor patterns confirmed that related mental abilities had been brought into play in the solution of the tests. The lack of significant loadings on the test of verbal intelligence suggested that these abilities were certainly non-verbal in nature. This suggestion was corroborated by the somewhat higher loadings found on the non-verbal test at each year-level. Having further regard to the actual material in which the ability tests were couched (arithmetical numbers; algebraic symbols; spatial figures) it seems likely that the abilities brought into play could be said to be specifically mathematical in nature, rather than generally non-verbal. Under this hypothesis, these specifically mathematical abilities form a sub-group in the

total field of non-verbal abilities. Turning now to the interpretation of the actual results obtained at the first, second and fourth year-levels, the two major factors were identified as mathematical factors of 'variable-class' and 'order-correspondence' types. At the third and fifth year-levels, the pattern of factors was not so well defined, though evidence of similar factors could be seen; these patterns were probably marred by some unsuitable non-normal distributions. At each year-level a small third factor was extracted, and the nature of this factor varied in each case. These third factors carried only a small number of significant loadings, and were identified as factors of 'variable-correspondence,' 'order,' 'verbal,' 'variable-order' and 'numerical variable-order' types at the first, second, third, fourth and fifth year-levels respectively. Four factors, corresponding to the four abilities of 'recognition of variable,' 'classification,' 'ordering' and 'recognising correspondences' might have been expected at each year-level. However, the joint psychological factors obtained in this work may be defensible from the stand-point of the philosophical inter-dependence of the concepts of variable and class, or of order and correspondence, discussed in Section II. It may be that such factors show the smallest break-up of abilities which can be expected in work of this kind. In any case, the present results represent an advance on previous work, in that the factors obtained are not obscured by other 'material' factors associated with number and space. In addition, the major factor pattern is stable throughout the successive year-levels. The extent to which these joint factors were found to account for the correlations between the ability and attainment tests suggests that the abilities called into play in the solution of the tests of mathematical ability also contributed to the solution of the attainment tests. Some further abilities, no doubt of a general or group nature, are also required for the successful completion of the latter tests. On the other hand, it seems likely that these two batteries of ability and attainment tests in Mathematics share a common ground in that success in the former seems to predict success in the latter, in the sense in which ability and attainment in the study of a subject were defined at the commencement of this work.

VII.—SUMMARY.

1.—Twelve tests of mathematical ability and fifty-three attainment tests were constructed in an attempt to assess how far mathematical ability enters into success in school mathematical work. Two intelligence tests (one verbal and one non-verbal) were combined with the ability tests.

2.—The intelligence tests and the tests of mathematical ability were administered at the end of the Summer term to one-hundred boys and girls in each of the five-year levels of a grammar school course. The attainment tests administered to children at any one year-level were based on the corresponding mathematical work studied by the group in the course of the year.

3.—Examination of the raw test scores showed that no testee obtained consistently low marks for the tests of mathematical ability while obtaining consistently high marks for the corresponding attainment tests. On the other hand, quite a number of testees scored well on the ability tests and badly on the attainment tests.

4.—Few significant differences were found between the boys' and girls' performances in the tests.

5.—Factor analyses of the correlations between the various intelligence and ability tests confirmed the interplay of mental abilities corresponding to the

subjective structure of mathematical thinking upon which the tests of mathematical ability had been founded, but did not reveal single factors corresponding to the 'stages' of mathematical reasoning.

7.—An extension of the factor analysis technique showed that the mental abilities involved in the working of the tests of mathematical ability also entered into the working of the attainment tests.

8.—Correlations ranging from .13 to .57 were obtained between the tests of mathematical ability and the School Certificate Mathematics marks at the fifth year-level. Correlations between the tests of mathematical attainment and the School Certificate marks ranged from .45 to .74.

VIII.—APPENDIX OF TESTS OF MATHEMATICAL ABILITY.

(1) SAMPLES OF TESTS OF MATHEMATICAL ABILITY.

The instructions and opening questions of the arithmetical number tests are given below. Codes for the tests are used as in Table I of the article.

Variable-Number (V.N.).—In the following rows of figures, the six of them are each a different numerical form of a certain kind of number. Put down on the right-hand line what kind of number is represented in each case.

Example : $1^2, 6^2, 12^2, 3^2, 4^2, 17^2$ the square of a number.

Each of the numbers on the left is the square of a number. Hence, "the square of a number" is the kind of number represented in this case. Now Do THESE :

1. $\frac{2}{-2}, \frac{5}{5}, \frac{6}{6}, \frac{11}{11}, \frac{7}{7}, \frac{12}{12}$

2.—2, 6, 12, 22, 26, 14

etc. (20 questions).

Class-Number (Cl.N.).—In each of the rows of numbers, there are five numbers on the left. Four have something in common and one is different. Cross out the one which is different. When you have done this, underline one of the numbers on the right which could best take the place of the one you have crossed out.

Example : 8, ~~3~~, 4, 22, 18 7, 14, 31, 19

3 has been crossed out because it is odd, and 14 is underlined because it is the only one of the four numbers on the right which is even. Now Do THESE :

1.—15, 25, 44, 45, 75 19, 15, 16, 24

2.—45, 54, 63, 27, 30 36, 29, 93, 15

etc. (20 questions).

Order-Number (O.N.). The following rows show the beginnings of series of numbers, and some numbers are missing. Fill in the blank spaces with the numbers you think should be there.

Example : 2, 4, 6, 8, 10, 12, 14

In the given series of numbers, each one is obtained by adding 2 to the previous one. Hence 12 and 14 fill the blank spaces. Now Do THESE :

1. 1, 3, 5, 7, 9, —, —,

2.—1, $\frac{1}{11}, \frac{1}{21}, \frac{1}{31}, \frac{1}{51}, \frac{1}{71}$

etc. (20 questions).

Correspondence-Number (C.N.) (Part I).—In the following rows of figures, all those in the Y line are obtained from the corresponding figures in the X line by performing the same operation. See if you can find how it is done, and then fill in the blank spaces. Try and write the calculation on the dotted line, using X and Y.

Example :

X	5	9	1	4	3	6	8	7	
Y	10	18	2	8	6	12	16	14	

$$Y=2X.$$

The blank spaces (now underlined) were filled by 18 and 16 because each of the other numbers in the Y line are twice the corresponding number in the X line. The calculation is $Y=2X$, because *any* Y is twice the corresponding X. Now Do THESE :

1.—X	5	2	7	3	1	8	19	11	101	4
Y	25		49	9		64	361	121	10201	16

etc. (5 questions)

Correspondence-Number (C.N.) (Part II).—Extension to similar relationships between three figures. Typical question to be completed for the number Z in terms of X and Y :

1.—X	7	9	3	2	12	11	4	1	0	10
Y	3	2	7	1	5	8	4	10	6	12
Z	21	18	21	2	60		16	10		120

etc. (5 questions).

To each of these four tests were added two exactly parallel tests couched in algebraic symbols and spatial figures instead of arithmetical numbers. The form of test remains the same. The tests are thus twelve in all.

IX.—ACKNOWLEDGMENTS.

The writer is indebted to the Central Research Fund, University of London, for the substantial financial help which rendered this investigation possible. She would also like to express her gratitude and thanks to the many people who helped in other ways.

X.—REFERENCES.

- ¹ BROWN, W. (1916) : *An Objective Study of Mathematical Intelligence*, D.Sc. Thesis, University of London Library.
- OLDHAM, H. W. (1937) : "A Psychological Study of Mathematical Ability with Special Reference to School Mathematics, Part I," *Brit. J. Educ. Psych.*, VII, p. 269.
Do., Part II, *Brit. J. Educ. Psych.*, VIII, p. 16.
- ² HAMLEY, H. R. (1934) : *Relational and Functional Thinking in Mathematics*. (Ninth Yearbook, National Council of Teachers of Mathematics, Bureau of Publications, Teachers' College, Columbia University, New York City, U.S.A.)
- ³ BENNETT, W. (1948) : *An Investigation into Mathematical Abilities Most Closely Related to Logical and Critical Thinking*. (M.A. Thesis, University of London Library.)
- BLACKWELL, A. M. (1940) : "A Comparative Investigation into the Factors Involved in Mathematical Ability of Boys and Girls," Parts I and II, *Brit. J. Educ. Psych.*, X, p. 143 and p. 212.
- MITCHELL, F. W. (1937) : *The Nature of Mathematical Thinking*. (Ph.D. Thesis, University of London Library.)

- JENKINS, J. W. (1939): *An Analysis of Factors Entering into the Results of Tests Based on the Logical Principles of Mathematics*. (Ph.D. Thesis, University of London Library.)
- HAMZA, M. (1952): "Retardation in Mathematics Amongst Grammar School Pupils," *Brit. J. Educ. Psych.*, XXII, p. 190, Table I.
- ⁴ HAMLEY, H. R., *loc. cit.*, p. 194.
- ⁵ HAMLEY, H. R., *loc. cit.*, p. 203.
- ⁶ JENKINS, J. W., *loc. cit.*, Chap. II.
- ⁷ See previous paragraph.
- ⁸ RUSSELL, B. (1919): *Introduction to Mathematical Philosophy*. (George Allen and Unwin, Ltd., London.)
- ⁹ JENKINS, J. W., *loc. cit.*, Chap. III.
- ¹⁰ See, for example:
- BLACK, M. (1933): *The Nature of Mathematics—A Critical Survey*. (Kegan Paul, Trench, Trubner and Co., Ltd., London.)
- RAMSEY, F. P. (1931): *The Foundations of Mathematics and other Logical Essays*. (Kegan Paul, Trench, Trubner and Co., Ltd., London.)
- WHITEHEAD, A. N. (1933): *The Aims of Education and other Essays*. (Cambridge University Press.)
- ¹¹ LONG, J. A., and SANDIFORD, P. (1935): *The Validation of Test Items* (University of Toronto, Department of Educational Research, p. 10.)
- ¹² Dr. C. M. FLEMING, University of London Institute of Education.
- ¹³ For this method the writer is indebted to Professor P. E. Vernon, University of London Institute of Education.
- ¹⁴ EMMETT, W. G., and WILMUT, F. S. (1952): "The Prediction of School Certificate Performance in Specific Subjects," *Brit. J. Educ. Psych.*, XXII, p. 52.
- ¹⁵ PEEL, E. A. and RUTTER, D. (1951): "The Predictive Value of the Entrance Examination as Judged by the School Certificate Examination," *Brit. J. Educ. Psych.*, XXI, p. 30.

THE VALIDITY AND INTERCHANGEABILITY OF TERMAN-MERRILL AND MATRICES TEST DATA

By D. WALTON

(Clinical Psychologist, Rainhill Hospital, near Liverpool, and Winwick Hospital,
Warrington.)

I.—Introduction. II.—The problem. III.—Results. IV.—Observations upon
results. V.—Discussion. VI.—General comment. VII.—Summary and
conclusions. VIII.—References.

I.—INTRODUCTION.

THE progress of research in diagnostic test construction and procedure would now seem to have reached a point where there is sometimes a serious lag in the incorporation of available knowledge into the practical application of such tests.

Some reflection on this situation was prompted by results obtained in the course of routine intelligence testing at a Home Office Classifying School. The tests used were the Terman-Merrill, Form M and Matrices (1938). The paper will present these results and attempt to show how significant are the misconceptions to which they may give rise, when viewed only in terms of what appear to be certain currently prevailing methods of interpretation and procedure.

There follows now an account of the test procedure and the results obtained.

II.—PROBLEM.

In the course of routine intelligence testing, significant discrepancies between intelligence quotients obtained by the Terman-Merrill, Form M, and Progressive Matrices (1938) gradings, were observed. These differences were mostly in the direction of lower Matrices gradings. The Terman-Merrill was administered to each child by the resident educational psychologist while the Matrices Test was given in group form and supervised by members of the teaching staff at the school. An average of three children at a time were tested with the Matrices. It was found necessary to adopt the latter procedure as it was impossible for the educational psychologist to test each boy individually with both tests. The rate of entry to the school was too great.

The Matrices was given as a non-verbal equivalent to the Terman-Merrill, i.e., as a non-verbal test of intelligence.

This diagnostic psychological testing procedure would appear suspect in view of much that has been written about the Matrices Test (Raven 1938, 1948). The present paper, therefore, attempts to present evidence against this testing procedure and to discover by an examination of test data from 360 approved-school boys, possible reasons for the discrepancies, noted on routine testing, between the Terman-Merrill and Progressive Matrices.

III.—RESULTS.

The results were obtained from 360 intermediate approved-school boys. They were all between 13 and 15 years of age.

Table I gives the frequency distribution of the two tests when Terman I.Q. and Matrices gradings are employed. The table was constructed for comparison with Table VII in *Guide to Using the Progressive Matrices* (1938). In the latter table a clinic group of 301 children was given the Matrices ('38) and Terman-Merrill *individually*. The relationship between Percentile Grade and Terman-I.Q. was specified.

TABLE I

RELATIONSHIP BETWEEN TERMAN I.Q. AND MATRICES GRADINGS FOR 360 INTERMEDIATE APPROVED SCHOOL BOYS.

Matrices Grade	Terman I.Q.					Total	Per cent. Frequency Approved School Boys	Raven's Sample
	Under 73	Under 89	89-111	Over 111	Over 127			
I	—	—	3	2	—	5	1.39	9.63
II	—	—	2	7	4	13	3.61	18.60
III	4	22	58	11	2	97	26.94	43.85
IV	20	74	67	4	1	166	46.11	15.95
V	22	44	13	—	—	79	21.94	11.96
Total	46	140	143	24	7	360	% frequency	
	12.78	38.89	39.72	6.67	1.94	Approved School Boys		
	12.29	18.27	45.52	13.29	13.62	Raven's Sample		

IV.—OBSERVATIONS UPON RESULTS.

(1) There is a concentration of scores at the low end of the scale for the two tests. The Matrices, however, shows a greater concentration at the lower end than does the Terman-Merrill, i.e., the Terman-Merrill shows considerably more subjects reaching the higher levels.

(2) When administered individually, as in Raven's sample, the Matrices results approximate more closely to those obtained using the Terman-Merrill. When administered in group form large discrepancies occur between it and the Terman-Merrill. Whilst only 12.78 per cent. of the approved-school boys had I.Qs. under 73, as many as 21.94 per cent. obtained a group Matrices grading of V. This is in contrast to Raven's figures which show 12.29 per cent. with Terman I.Qs. under 73 and 11.96 per cent. with an individual Matrices grading of V. In fact, 68.05 per cent. of the approved-school boys had Matrices gradings of 4 or 5 compared with 51.67 per cent. with Terman-Merrill I.Qs. equal to or less than 89. In contrast, Raven's figures show 27.91 per cent. to have Matrices gradings of 4 or 5, and 30.56 per cent. to have I.Qs. equal to or less than 89.

V.—DISCUSSION.

Three principal factors appear to have contributed to the low Matrices results:

- The test was administered under group conditions. Under these conditions children seem less likely to exert themselves, especially the dull, non-conscientious ones.
- The teachers were not fully qualified or trained to be responsible for psychological testing.
- The Matrices was used as a non-verbal equivalent of the Terman-Merrill. This equivalence seems to be an unjustified assumption. The Matrices is not intended to be used as a test of intelligence, equivalent to such a test as the Terman-Merrill, regardless of conditions.

It may be argued that the group testing of approved school boys with the Matrices, whilst desirable because of the time saved, does not seem justifiable from the results obtained. Boys with personality problems such as occur frequently in approved school populations, seem to react better to the individual test situation where they are the centre of attention rather than to the group test situation in which there is an impersonal element. This would seem particularly the case for dull and non-conscientious types of children. The result is that they do not trouble to exert themselves because it is fairly easy on the Matrices to choose a plausible answer. However, if the test is administered individually and with adequate encouragement given, the child is more likely to respond by giving of his best.

Nevertheless, it can be argued that in real life situations these children are not always going to be the centre of attention, be given encouragement and helped to keep their attention on what they are doing. In this sense then the low Matrices results may be more valid assessments, "... of their natural capacity (i.e., irrespective of such conditions as motivation) to develop a systematic method of reasoning than the Terman-Merrill Intelligence Quotient ultimately proves to be ... it is just this capacity not only to think clearly, but to develop a systematic method of reasoning, which plays the important part in acquiring knowledge and skill in any walk of life (Raven 1954)."

Often, however, we need a true assessment of their "... capacity to develop a systematic method of reasoning," irrespective of how they are likely to behave outside the school. To achieve this it is necessary to provide conditions similar to those recommended in the test instructions, such as allowing the child to work quietly at his own speed from the beginning to the end of the scale without interruption and under any further conditions which are suited to the individual needs of the child.

One further point needs to be emphasized. If one set of Terman-Merrill I.Qs. and one set of Matrices Gradings are available from the same sample, and if the Matrices Gradings are lower, contrary to the opinions which are often expressed, we are not justified in saying that the Matrices results are incorrect. It would be a mistake to suggest that even if the Matrices were administered under the best conditions it would ever give a correct appraisal of a subject's intelligence. Of the two tests only the Terman-Merrill is reputed to be a test of intelligence. Misinterpretation often results because inferences are drawn from the Matrices which are not justified by what the test is stated to do. Likewise I.Qs. are sometimes calculated from the Matrices scores and used as valid equivalents for Terman-Merrill I.Qs. However, as stated both in Raven's *The Comparative Assessment of Intellectual Ability* (1948) and *Guide to Using the Progressive Matrices* (1938), "Progressive Matrices ... is not a test of 'general intelligence,' and it is always a mistake to describe it as such."

The possibility of administrative errors by non-psychologically qualified teachers cannot be passed over lightly. It calls for a thorough training in scientific method and test construction for those administering psychological tests. In this way a greater appreciation of standardized test instructions, administration and procedure would be acquired.

With regard to evidence against the present testing procedure, it is admitted that whilst those debarred from adequate cultural stimulation are likely to be penalised by a heavily weighted verbal intelligence test, and that a non-verbal intelligence test would assess more accurately their innate ability or capacity, we must not think that we possess these ideal measuring instruments. Before a test can be given as a non-verbal equivalent it must be both shown that the test measures the same abilities as the other test, though its content is different, and that the results can readily be compared on the basis of a common score.

Although Burt (1954) suggests that Binet attempted to do what factorists nowadays attempt by more vigorous methods, i.e., to take those functions which appear most highly weighted with the general factor at each successive age, and then measure that factor by averaging the sum of the obtained scores, this does not appear to have been achieved. McNemar (1942) found that the scores were heavily weighted by abilities other than general intelligence, in fact for the average Terman-Merrill item, the general factor accounted for about half of what the item measured. The remaining variation was accounted for by miscellaneous knowledges and abilities. Group factors were rarely found in more than one or two items, consequently it was not possible to identify these factors as separate scores. The Matrices ('38) on the other hand has been shown to have a G saturation varying between 0.8 with negligible loadings of other factors (Slater 1948) and 0.82 (Burt). Two further points are important. First, the Terman-Merrill represents somewhat different abilities at different ages. Secondly, it differs from the Matrices in that it contains both untimed and speed tests. In other words, the two tests consist of combinations of different abilities in different proportions and demand somewhat different mental operations. On the basis of these observations there appears little justification for using one as a non-verbal equivalent of the other.

VI.—GENERAL COMMENT.

It is known that Matrices results are sometimes converted into equivalent I.Qs. and compared with the Terman-Merrill. This is unjustified. The intellectual capacity as measured by the Matrices test does not develop uniformly throughout childhood, consequently one is not justified in calculating I.Qs. from the test scores and comparing them with Terman-Merrill I.Qs.

To be able to calculate I.Qs. one basic assumption must be fulfilled. Intellectual development must be uniformly proportional to chronological age. There is, however, increasing evidence against the existence of a linear relationship between them. Raven⁷ (1948) reached the following conclusion about the interpretation of the Matrices (1938) and Mill Hill (1943) in this light, "It is clear . . . that the two intellectual processes essential for success in a test of general intelligence do not develop at the same rate, do not mature at the same age, do not remain constant for the same length of time, or decline at the same rate. The assumptions involved in calculating intelligence quotients . . . are therefore, untenable."

Wechsler (1944) found similar results to Raven. He constructed a logistic curve of mental growth based on the results obtained with the Belle Vue Intelligence Scale. The age range was 7-17. It showed that intellectual growth, as measured by the 'Belle Vue,' did not advance by equal amounts throughout its development, ". . . any method of calculating I.Qs. which assumes a linear relationship between chronological and mental age cannot possibly give constant values for any considerable portion of the growth period." The results are that for the average individual the mean I.Q. will change from age to age. This is seen on the revised Stanford-Binet. The mean I.Qs. are above 100 for most ages. Marks (1944), for instance, has shown that all means except for ages 4 and 5½ are over 100 and seven age groups have average I.Qs. over 103. The second consequence is that a child's I.Q. is likely to "fall off" as he grew older. This would be more pronounced for people of low intelligence. This phenomena can be explained by an examination of the logarithmic relationship between mental and chronological age. The numerators increase more slowly than do the denominators. The difference is more accentuated in the case of mental defectives (Kuhlmann, 1921).

VII.—SUMMARY AND CONCLUSIONS.

1.—The present experiment was designed to discover reasons for large discrepancies noted on routine testing of approved-school boys, between the results of the Terman-Merrill and Progressive Matrices tests.

2.—The results from 360 intermediate approved-school boys, tested individually with the Terman-Merrill, Form M, and in groups with the Progressive Matrices ('38), were therefore compared with those from a clinic group of 301 children who had been tested individually with the Matrices ('38) and Terman-Merrill.

3.—When the two tests are administered individually, as in Raven's sample, the results show closer agreement than when the Matrices is administered in group form. The following reasons are suggested to account for this difference :
(a) It appears that children with personality problems react better to the individual test situation and consequently provide more valid results.

(b) Since it is fairly easy on the Matrices to choose a plausible answer, then dull, non-conscientious children are particularly liable to provide invalid results under group conditions of testing. It will be noted that Raven's sample contains a very much larger percentage of children with I.Qs. in the above average range than does the approved-school group.

These difficulties would appear to be increased when the administration and supervision is the responsibility of teachers who are not fully trained to be responsible for psychological testing.

4.—The principal implication of this investigation is to suggest the need for carefully controlled studies to determine the laws of interaction between disturbed personalities reacting under both individual and group conditions of testing. Level of intelligence seems to be an important variable to be taken into account in any such investigations.

VIII.—REFERENCES.

- ¹ BURT, C. (1954) : " The Differentiation of Intellectual Ability," *Brit. J. Educ. Psych.*, Vol. XXIV, Part II.
- ² KUHLMANN, F. (1921) : " Results of repeated Mental Re-Examination of 639 Feeble-Minded over a Period of Ten Years," *J. Appl. Psych.*, V, 192-224.
- ³ MARKS, E. S. (1947) : " Sampling Errors in the Revision of the Stanford-Binet," *Psych. Bull.*, 44, 413.
- ⁴ MCNEMAR, Q. (1942) : *The Revision of the Stanford-Binet Scale*. (Boston : Houghton Mifflin.)
- ⁵ RAVEN, J. C. (1938) : *Progressive Matrices*. (London : H. K. Lewis and Co., Ltd.)
- ⁶ RAVEN, J. C. (1943) : *The Mill Hill Vocabulary Scale*. (London : H. K. Lewis and Co., Ltd.)
- ⁷ RAVEN, J. C. (1948) : " The Comparative Assessment of Intellectual Ability," *Brit. J. of Psych.*, Vol. XXXIX, Part I.
- ⁸ RAVEN, J. C. (1948) : *Guide to Using Progressive Matrices*. (London : H. K. Lewis and Co., Ltd.)
- ⁹ RAVEN, J. C. (1954) : Private Communication.
- ¹⁰ SLATER, P. (1948) : " Comment on ' The Comparative Assessment of Intellectual Ability,' " *Brit. J. of Psych.*, Vol. XXXIX, Part I.
- ¹¹ TERMAN, L. M., and MERRILL, M. A. (1937) : *Measuring Intelligence*. (London : Harrap.)
- ¹² WECHSLER, D. (1944) : *Measurement of Adult Intelligence*. (Baltimore : Williams and Wilkins.)

ACKNOWLEDGMENTS.

The writer wishes to thank Professor C. W. Valentine, Professor P. E. Vernon, and Mr. J. C. Raven for their valuable criticisms and suggestions given during the preparation of this paper for publication.

A COMPARATIVE STUDY OF EDUCATIONAL ATTAINMENTS IN ENGLAND AND SCOTLAND

By P. E. VERNON

(*University of London, Institute of Education*)

AND

M. B. O'GORMAN and A. McLELLAN

(*University of Glasgow, Psychology Department*)

I.—*Introduction.* II.—*The samples.* III.—*The tests.* IV.—*Results.*
V.—*Summary.* VI.—*Appendix.* VII.—*References.*

I.—INTRODUCTION.

It is generally agreed that Scottish primary education tends to be more formal than English, involving the earlier introduction of, and more thorough drilling in, the three R's, and a greater use of corporal punishment. Writing in 1938, Vernon referred in his *Measurement of Abilities* to "the appalling efficiency of Scottish teaching methods" (which) "raises the performance of Scottish elementary school children on tests of the three R's by anything from 10 per cent. to 50 per cent. over the English level." And McGregor found, in 1934, that 11-year-old Fife children were roughly two years superior to the American norms on the arithmetic, language and spelling sections of the Public School Achievement Tests, though only thirteen months above on Reasoning Arithmetic and 5 months on Reading Comprehension. As evidence of the supposed further superiority of Scottish secondary education, the larger numbers of pupils entering senior secondary schools than of English pupils entering grammar schools, and the much higher proportion of graduate teachers, are often instanced. However, on the basis of work with a graded word reading test, Vernon (1938) suggested that English educational standards may catch up with Scottish ones after the transfer to secondary schools. And his survey of the abilities of 9,183 18-year Army recruits in 1947 yielded a mean Educational Quotient for Scottish recruits of 98.2, as against 100.0 for Welsh and 100.2 for English; though relative to their g-test scores, the Scots were still superior educationally.

Now, quite apart from national prestige, it is of the utmost educational importance to determine the facts more precisely. Do Scottish methods (in so far as they show any general difference from English) actually produce better results, and is their effect a lasting one? Or do they, at the cost of greater pressure on young children, produce a merely temporary advantage and possibly even a relative decline when pupils reach adolescence and become more rebellious to adult authority? A proper answer to such questions would require a large-scale survey by such organisations as the National Foundation for Educational Research in England and Wales, and the Scottish Council for Research in Education. The Director of the National Foundation (Morris, 1955) has described a similar project in France and Belgium where "the hypothesis to be tested is whether "the initial superiority of French children "is maintained up to the end of the primary school or whether it gradually decreases and perhaps vanishes, or is even reversed"; and he recommends co-operative studies in Britain. The present investigation represents a very small-scale contribution carried out by Vernon in England, and by O'Gorman and McLellan

in Scotland. It was confined to the schools in one English and one Scottish town and is, therefore, obviously open to the criticism of being unrepresentative of the countries as a whole. But there is also an obvious answer to those who raise this criticism, namely that they should carry out further similar studies. The writers would be glad to make their tests available for such a purpose. Unrepresentativeness was also guarded against to some extent: (a) by covering almost the complete school populations of the towns concerned, i.e., six primary and four secondary schools in England and four primary and four secondary schools in Scotland; (b) by the inclusion of intelligence tests known to have low verbal content, in order to equate the populations more exactly.

II.—THE SAMPLES.

The two towns were carefully chosen to be comparable in size and in occupational distribution. E- is a country market town in the Home counties, with some 14,000 inhabitants. S- is a seaside town of about 10,000. Both have a little light industry and a considerable element of commuting business men.

It was decided to test three age groups, namely 7:6 - 8:5, 10:9 - 11:8 and 14:0 - 14:11. These are referred to below as the 8, 11 and 14-year-olds. All children (apart from a few absentees) in these age ranges in the Local Authority schools, were included, regardless of their classes. Almost all the 8's were in their second or third term in the bottom junior class. The 11's raised certain difficulties, since the peak of effort in English and Scottish primary schools occurs at different ages. English pupils normally take the selection examination at 10½-11½, while the Scottish ones mostly take the Qualifying or Control at 11½-12½. It seemed that the fairest plan was to give the tests in June-July, so that the E- pupils would have had some three months' relative relaxation after their peak, while the S- pupils would be working up to their's some eight months ahead.* As regards the 14's, just one-quarter of the E- pupils get into the local grammar schools (which also take children from surrounding areas, who were excluded from the testing), and three-quarters into modern schools. In S- also, one-quarter enter senior secondary schools, but another 12 per cent. take a two-language course in a high school and transfer at 15 to complete their full secondary course. These 12 per cent. were among the brightest in the age group, hence they were classified with the senior secondary group.

In neither town do appreciable numbers transfer to secondary schools in other areas. At all ages there are 5-10 per cent. in private, preparatory or public schools, who could not be tested. Enquiries indicated that this is a fairly constant element. Between 11 and 14 a few children leave the private and enter the State schools, and about an equal number do the reverse. But the possibility of differential loss could not be ruled out, and constitutes another legitimate criticism of this limited survey.

The numbers of testees are shown in Table I.

III.—THE TESTS.

The tests were mostly 'tailored' to suit the populations, though based on well-established instruments. All had been previously tried out and item-analysed, and were provided with appropriate practice material and instructions designed to ensure identical application in the two towns. The fact that they were not standardised did not matter, since norms were derived from the

* Sutherland (1951) has shown an average loss in scores on Moray House arithmetic tests of 2 to 3 points, a year after sitting the Scottish transfer examination, and attributes this largely to the drop in pressure on the pupils.

TABLE I
NUMBERS TESTED IN EACH AGE GROUP.

Age Group	English		Scottish	
	Boys	Girls	Boys	Girls
Eight	103	80*	54	62
Eleven.....	85	84	64	74
Fourteen—grammar and senior	26	17	21	19
Fourteen—modern and junior	58	66	35	33

groups actually tested. Distributions were tabulated for equal numbers of English boys, English girls, Scottish boys and Scottish girls, and then combined. The median M.A. or E.A. was now equated to the median C.A. of the whole group, and higher and lower scores were converted via percentiles to yield I.Qs or E.Qs. with a uniform standard deviation of 15.

Eight-Year.—(1) Moray House Picture Intelligence Test. Time did not allow the application of the whole of this to all children, and sub-tests 2, 3, 4, 5 and 6 (lasting some 45 minutes in all) were selected. However, sixty-four pupils did the complete test, and a correlation of ± 0.944 was obtained between the abbreviated and full-test I.Qs. The published norms (at this age level) do not seem quite accurate, since the median I.Q. of the combined groups was 110, although every effort was made to adhere to the proper instructions. Our two towns may be a little above the average of the general population, but certainly not 10 points. Hence the abbreviated scores were re-scaled by the method already indicated.

(2) Word Recognition: 11 items. The test blanks contain rows of five words, e.g., give, good, gave, glad, grave, and the tester reads out one word in each row, which the pupils encircle.

(3) Reading Comprehension: 13 items requiring the writing of one word in a blank space, e.g., We wear shoes on our
The time limit was 8 minutes.

(4) Spelling: 12 words dictated from Burt's list.

(5) Arithmetic: 20 four-rules sums (15 minutes).

Fuller details are given by O'Gorman (1954).

Eleven-year.—The first five tests were the same as those used in Kemp's (1955) survey of London primary schools, namely:

(1) Abstraction test of intelligence (22 items, 10 minutes).

(2) Reading comprehension (25 items, 6 minutes).

(3) Spelling (21 words).

(4) Mechanical arithmetic (20 items, 7 minutes).

(5) Problem arithmetic (22 items, 10 minutes).

It was desired to include a General Knowledge test and an English essay or composition; but to give both of these after the above battery would have been too fatiguing for the pupils. Thus, in each school the pupils were split into half-groups, whose equivalence was confirmed by their Reading Comprehension test scores.

(6) General Knowledge: 50 multiple-choice items from the fields of geography, science, art, literature, music, sport and current affairs; no time limit. This was constructed by McLellan, but no bias could be detected towards

* No reason could be found for this irregularity in numbers of boys and girls. But it is quite insignificant statistically, chi-square being 2.9.

knowledge more likely to be familiar to Scottish than to English pupils, or the reverse.

(7) Composition on the title 'Myself.' No hints were given, and 30 minutes were allowed. The same sheets of paper were used in all schools, and names were written on the back. All sheets were then shuffled together and marked on a letter-scale independently by one Scottish and one English marker, who did their best to disregard any indications of country of origin. The letter-grades were converted to standard scores, and a correlation of 0.799 was obtained between the markers. Fortunately, there was no tendency for the markers to disagree in their averages for the two countries, and their marks were simply added to yield a final distribution. It is relevant to note that essays are included (with other English tests) in the transfer examinations in both areas.

Fourteen Years.—Most of the tests were extended versions of those used with 11-year-olds.

(1) Abstraction (15 minutes).

(2) Reading Comprehension (10 minutes).

(3) Spelling: an objective test was substituted consisting of thirty sentences, each with one mis-spelt word to be re-written (10 minutes).

(4) Arithmetic-Mathematics: 60 short sums or problems, similar to those in the Vernon Graded Arithmetic-Mathematics test, ranging from four rules to elementary algebra and geometry (20 minutes).

(5) General Knowledge (half-group, no time limit).

(6) Composition on 'Myself' (half-group). The same procedure was followed for marking. Although the correlation between markers was 0.783, there did appear at this level a strong tendency for one or both markers to favour his own national style and content. A third marker (not one of the authors) was enlisted, namely a Scottish-educated and trained teacher who had taught chiefly in English secondary schools. Her relative assessments differed again, and it may be that too much depends on the personal inclinations of the marker for any clear consensus of opinion to be reached. Under the circumstances it seemed better simply to average the first two markers, despite their discrepancies, rather than attempt to find a wholly unbiased set of marks.

It was not possible to include a test of French or other advanced secondary subjects, since no common basis could be found in the curricula of the two levels of schools.

IV.—RESULTS.

Eight Years.—The full distributions would take up a great deal of space and would be difficult to interpret. We shall, therefore, quote the means for boys and girls in each town, and list in the Appendix abbreviated distributions giving the numbers in each group obtaining scores in the top 15 per cent, the next 35 per cent., the next 35 per cent. and the bottom 15 per cent.—referred to below as A, B, C and D scores. The latter will indicate the extent to which any group contains a large 'tail' of D-scorers, or a large proportion of high A-scorers.

The groups are less closely equivalent in intelligence than had been hoped.* Probably E- contains more professional and business parents than S-. Thus, any educational superiority found among the Scottish pupils should be larger still for children of the same I.Q. It will be seen that the Scottish children are greatly advanced in formal subjects, though not in reading comprehension. By 8 years all but a very few of the dullest have begun to read, whereas English

* At the same time, this difference is almost the same as that found between all Scottish and English recruits on *g*-tests (Vernon, 1951). Thus, it may be reasonable to interpret the educational test results as typical of the two countries also.

TABLE II
MEAN QUOTIENTS AND DIFFERENCES* FOR EIGHT-YEAR-OLDS.

	English		Scottish		Mean Scottish superi- ority	Mean Girls' superi- ority
	Boys	Girls	Boys	Girls		
Intelligence (picture test) .	102.7	101.6	99.3	96.4	+3.3	-2.0
Word Recognition	96.6	98.3	102.3	102.7	+5.0	+1.0
Reading Comprehension . .	98.6	101.7	99.3	100.4	-0.3	+2.1
Spelling	94.2	96.9	104.0	104.9	+8.9	+1.8
Arithmetic	93.8	96.0	108.6	101.6	+10.2	-2.4

schools are content to carry a long tail of non-readers. In the Appendix it is shown that, on the three verbal tests, there are 26 per cent. of D-scorers among English boys, 8 per cent. among Scottish. The girls are somewhat less divergent. But English schools perhaps provide better opportunities for brighter juniors to progress, there being 18½ vs. 11 per cent. of A-readers on the comprehension test. In arithmetic, Scottish boys at all levels are far ahead, easily surpassing the girls; whereas English boys are behind their sisters.

The effectiveness of the Scottish system casts considerable doubt on many of the claims made by American psychologists for minimum maturational ages before introducing the formal subjects. It must be admitted that early acquaintance with the tool subjects has great advantages, and it should allow the Scottish pupil to progress more rapidly throughout his school career. Whether this advantage is obtained at the cost of psychologically undesirable strain among the infants, no-one can say. We would not for a moment accuse the teachers in S-, or elsewhere in Scotland, of being inhumane. They are following the system in which they themselves were brought up, and which they believe to be necessary. Nevertheless, the strap is certainly used in some of the S- schools, whereas corporal punishment is practically restricted to severe disciplinary offences, and given only by the head teacher, in E- schools. A far more comprehensive study would be needed to bring out the full psychological implications of the contrasting practices. Here we can only try to follow up the progress in attainments.

TABLE III
RESULTS FOR ELEVEN-YEAR-OLDS.

	English		Scottish		Mean Scottish superi- ority	Mean Girls' superi- ority
	Boys	Girls	Boys	Girls		
Intelligence	101.6	100.5	97.2	100.7	-2.1	+1.2
Reading Comprehension . .	101.1	98.9	98.3	101.7	0.0	+0.6
Spelling	94.8	96.0	100.7	108.5	+9.2	+4.5
Mechanical Arithmetic . .	96.6	96.9	101.3	105.3	+6.6	+2.1
Problem Arithmetic	100.9	98.3	99.1	101.6	+0.7	0.0
General Knowledge	103.2	99.4	100.9	96.5	-2.6	-4.1
Composition	95.6	101.8	99.9	102.7	+2.6	+4.5

* The S.E.s. of the differences in Tables II-IV (for total groups) are all close to 1.8. Thus, differences of ± 3.6 are conventionally significant at the .05 level, and smaller ones must be interpreted with particular caution.

Eleven Years.—With a slightly more verbally-loaded intelligence test, the national difference decreases and the sex difference is reversed. The educational tests show the same trend as before but, except in the case of spelling, they are decidedly smaller, and are only of the same order as the sex differences. The differences are also very similar to those reported by MacGregor between Scottish and American children, though much smaller. Thus, not only in reading comprehension, but also in reasoning arithmetic, there is little or no superiority (unless allowance is made for the lower Scottish I.Q. level). The General Knowledge and composition differences, being based on fewer cases, are too small to be statistically reliable, though the sex differences are probably meaningful, and are confirmed among 14-year-olds. Indeed, there is a large amount of evidence from other sources that girls and women are superior in spelling and writing, and inferior in general knowledge of most kinds. The Appendix further shows that there are more English than Scottish A-scorers among boys for composition, though also more D's. Perhaps the most striking features of the tables in the Appendix are the 36 per cent. English and 1 per cent. Scottish D-scorers among girls in Mechanical Arithmetic, and the 30 per cent. English boys vs. 2 per cent. Scottish girls scoring D in spelling.

Fourteen Years.—Table IV gives the results for all boys and girls, as before, and Table V lists the separate figures for different types of school.

TABLE IV
RESULTS FOR FOURTEEN-YEAR-OLDS.

	English		Scottish		Mean Scottish superi- ority	Mean Girls' superi- ority
	Boys	Girls	Boys	Girls		
Intelligence	100.6	100.6	98.8	100.1	-1.2	+0.6
Reading Comprehension ..	101.6	96.0	101.9	100.5	+2.4	-3.5
Spelling	94.5	98.2	99.7	107.7	+7.3	+5.8
Arithmetic-Mathematics ..	99.4	97.2	102.8	100.7	+3.4	-2.1
General Knowledge	103.2	94.1	103.0	99.6	+2.7	-6.2
Composition	98.6	101.7	93.8	106.0	-0.2	-7.6

Again the Scottish pupils are slightly inferior on the Abstraction test of intelligence, and are now markedly superior only in spelling. They seem to be a little better in reading, arithmetic-mathematics and general knowledge, though hardly to a statistically significant extent. The essay results, as already pointed out,* are doubtful, but it would seem that even the slight Scottish superiority found at 11 has been lost by 14. Thus, the national differences have not wholly disappeared, as suggested in the Introduction, but they are now much less marked than the sex differences.

The English modern and Scottish junior secondary schools show just the same pattern of differences as the total groups, except that these average and duller girls are relatively more superior in essay-writing and less so in spelling than are all girls. But the selective school results are rather different, largely because the E-grammar schools contain a smaller and more highly selected and, therefore, more intelligent population than the S-senior secondary schools. The English pupils are now superior in all subjects except for a slight deficit

* According to Vernon's scoring the English are 4.8 E.Q. points superior to the Scottish, according to McLellan's 2.0 points inferior.

in spelling, and particularly so at essay writing. This is brought out more clearly by the distributions in the Appendix, where it will be seen that there are very few, or no, C or D-scorers in the grammar schools, but an appreciable number on every test in the senior secondary schools.

In order to obtain an overall picture, the E.Qs. at different ages were combined as follows:

Eight Years—reading comprehension, spelling, arithmetic, equally weighted;

Eleven Years—reading comprehension, spelling, general knowledge, essay, mechanical arithmetic, problem arithmetic, equally weighted;

Fourteen Years—the same four verbal subjects, with arithmetic-mathematics doubly weighted.

TABLE V
RESULTS FOR FOURTEEN-YEAR-OLDS IN DIFFERENT LEVELS OF SCHOOLS.

	English				Scottish				National Differences*		Sex Differences*	
	Boys		Girls		Boys		Girls		Sen.	Jun.	Sen.	Jun.
	Gr.	Mod.	Gr.	Mod.	Sen.	Jun.	Sen.	Jun.				
Intelligence	113.7	93.1	115.1	95.7	105.7	91.1	109.8	94.5	-6.6	-1.6	+2.7	+3.0
Reading Comprehension	115.6	96.2	109.6	91.4	110.0	97.3	109.1	95.4	-3.0	+2.5	-3.4	-3.3
Spelling	110.1	89.3	105.7	95.7	108.7	94.4	113.7	104.1	+3.3	+6.7	+0.3	+8.0
Arithmetic-Mathematics	121.5	92.0	112.9	92.0	110.1	99.8	112.0	94.0	-6.1	+4.9	-3.3	-2.9
General Knowledge	120.8	97.2	106.3	90.1	111.1	98.3	111.4	92.6	-2.3	+1.8	-7.1	-6.4
Composition	115.4	91.2	116.1	98.2	105.9	87.7	114.6	100.7	-5.5	-0.5	+4.7	+10.0

Calculated in this way the average superiority of Scottish girls (disregarding intelligence level) is 4.1, 4.1 and 5.1 at 8, 11 and 14 years. Indeed, the mean E.Qs. and the distributions for girls on the combined tests remain remarkably similar throughout. With their greater docility than that of boys, they absorb a little more from the Scottish system and retain their lead. But for boys the corresponding figures are 8.5, 1.3 and 1.3. Apparently S-boys are forced much more rapidly than E- at first, but even by 11 years have begun to rebel and show very little difference from then on, and little benefit from their early advantage. Nevertheless, it would be unfair to conclude that there is no gain at all, in view of the somewhat lower intelligence test scores with which they started. Finally, we must point out again that not only are the numbers of schools tested in the two countries too limited to justify any but the most tentative generalisations, but also the numbers are too small within these schools (particularly when the total groups are sub-divided) for the detailed figures to be highly stable.

V.—SUMMARY.

1.—A comparative survey was made by English and Scottish psychologists of the educational attainments of all children in the Local Authority schools in an English and a Scottish town, at three age levels, 8, 11 and 14+.

2.—Though 881 children in eighteen schools were given tests covering the main school subjects, the data are much too limited for most of the obtained differences to be highly reliable; and there are further obvious dangers in generalising from these schools to Scottish and English education in general.

* These figures for grammar-senior vs. modern-junior do not appear to accord with the averages in Table IV. But this arises purely because of the unequal proportions in the upper-level schools in the two countries. The averages are quite correct.

3.—Despite the superficial comparability of the two towns, the English pupils were a little superior on group intelligence tests, hence any Scottish superiority in attainments is likely to be under-estimated.

4.—The Scottish 8-year-olds were greatly advanced in formal subjects—reading (word recognition), spelling and arithmetic, though not in reading comprehension. By 11 years there is little difference in reading, problem arithmetic, general knowledge and composition, though the superiority in spelling and mechanical arithmetic persists. The picture is generally similar at 14, but English grammar school pupils (mainly because of their higher degree of selection) are superior on most tests to Scottish senior secondary school pupils.

5.—With increasing age, sex differences become more marked than national ones. In general, Scottish girls retain much the same overall advancement throughout. But Scottish boys are forced far ahead by 8, and then lose almost all this advantage by 11.

6.—These results suggest that children can make progress with the three R's at an earlier age than some psychologists believe to be desirable. They certainly cannot show that this system is, or is not, psychologically harmful; but they do indicate that the lastingness of its educational benefits (at least among boys) is disappointing.

The author's most grateful thanks are due to the Directors of Education of the two areas who gave permission for this investigation, and to the headteachers and staff of the schools for their co-operation in the testing arrangements. Acknowledgments are due also to the Institute of Education, University of London, for financial assistance covering the cost of the printed and duplicated test materials and travelling expenses.

VI.—APPENDIX.

The proportions listed in the following tables were read off from percentage ogives. A = the top 15 per cent. of all scores on any test, B = the next 35 per cent., C the 35 per cent. below the median, and D the bottom 15 per cent.

E = English. S = Scottish. B = Boys. G = Girls.
g = grammar. m = modern. s = senior. j = junior.

8-Year Distributions.

Moray House					Word Recog.				Rdg. Compre.				Spelling			
EB	EG	SB	SG		EB	EG	SB	SG	EB	EG	SB	SG	EB	EG	SB	SG
A 22	13	11	15	}	42	45	54	59	18	19	7	15	10	13	17	20
B 34	43	34	29		32	38	40	31	32	40	36	33	22	24	49	44
C 35	31	40	33		26	17	6	10	25	25	47	42	40	44	27	29
D 9	13	15	23						25	16	10	10	28	19	7	7

8-Year

Arithmetic

EB	EG	SB	SG
7	9	30	14
25	31	47	37
41	38	19	41
27	22	4	8

11-Year Distributions.

Abstraction

EB	EG	SB	SG
19	15	7	19
36	35	33	34
31	36	43	31
14	14	17	16

Rdg. Compr.

EB	EG	SB	SG
21	15	9	16
33	33	34	40
28	37	42	33
18	15	15	11

Spelling

EB	EG	SB	SG
8	7	13	33
25	25	43	48
37	53	31	17
30	15	13	2

11-Year

Mech. Arith.

EB	EG	SB	SG
14	4	15	26
28	17	45	49
39	43	35	24
19	36	5	1

Prob. Arith.

EB	EG	SB	SG
21	12	11	16
33	31	37	37
27	41	37	37
19	16	15	9

Gen. Knowl.

EB	EG	SB	SG
24	13	12	10
34	33	45	28
29	42	29	41
13	12	14	21

Composition

EB	EG	SB	SG
17	11	12	19
29	29	41	41
33	42	34	31
21	18	13	9

14-Year : Abstraction

EB		EG		SB		SG	
g	m	g	m	s	j	s	j
14	—	16	6	8	—	11	4
8	28	9	25	19	15	20	16
3	31	—	28	10	34	6	30
—	16	—	16	—	14	—	13

Reading Comprehension

EB		EG		SB		SG	
g	m	g	m	s	j	s	j
19	3	8	3	9	6	11	—
6	30	10	18	21	17	18	22
—	27	7	30	7	30	8	30
—	15	—	24	—	10	—	11

Spelling

EB		EG		SB		SG	
g	m	g	m	g	m	g	m
8	2	2	2	10	—	19	15
14	12	17	20	20	17	16	25
3	33	6	40	7	33	2	17
—	28	—	13	—	13	—	6

Arithmetic-Mathematics

EB		EG		SB		SG	
g	m	g	m	s	j	s	j
20	—	9	—	16	—	15	—
5	19	16	19	15	30	17	20
—	34	—	38	6	26	5	31
—	22	—	18	—	7	—	12

General Knowledge

EB		EG		SB		SG	
g	m	g	m	s	j	s	j
19	1	4	—	15	6	15	—
6	34	15	16	14	15	18	21
—	27	5	31	7	41	4	24
—	13	1	28	1	1	—	18

Composition

EB		EG		SB		SG	
g	m	g	m	s	j	s	j
9	1	13	4	8	1	17	6
13	17	11	32	15	8	17	29
3	41	1	29	12	30	3	20
—	16	—	10	2	24	—	8

VII.—REFERENCES.

- KEMP, L. D. C. (1955): "Environmental and Other Characteristics Determining Attainment in Primary Schools," *Brit. J. Educ. Psych.*, 25, 67-77.
- MACGREGOR, G. (1934): *Achievement Tests in the Primary School*. (University of London Press.)
- MORRIS, B. S. (1955): "Some Impressions of Educational Research in Belgium and France," *Bull. Nat. Found. Educ. Res.*, No. 5, 19-25.
- O'GORMAN, M. B. (1954): *A Comparison of the Standards of Attainment in Certain School Subjects of Children in the Same Age Group in England and Scotland*. (University of Glasgow, Ed.B. Thesis.)
- SUTHERLAND, J. (1951): "A Comparison of Pupils' Arithmetical Ability in the Secondary School with their Ability at the Time of their Transfer from Primary Schools," *Brit. J. Educ. Psych.*, 21, 3-8.
- VERNON, P. E. (1938): *The Standardization of a Graded Word Reading Test*. (University of London Press.)
- VERNON, P. E. (1940): *The Measurement of Abilities*. (University of London Press.)
- VERNON, P. E. (1951): "Recent Investigations of Intelligence and its Measurement," *Eugen. Rev.*, 43, 125-137.

CRITICAL NOTE ON

"Testing Geography at the Ordinary Level of the General Certificate of Education"—by JOHN C. DANIELS.

(This *Journal*, Vol. XXIV, Nov., 1954.)

By S. S. DUNN,
Australian Council for Educational Research.

Daniels' paper purports to examine which of three examination question papers, i.e., purely essay questions, the interpretation of maps, diagrams and statistics, and short 'objective' questions, is the most reliable and valid.

It is not made clear why the interpretation of maps, diagrams and statistics is included since this could at best be considered a test of one element of geographical ability while the objective and essay tests were presumably meant to be 'global' tests. This test will not be further discussed except in so far as it is mentioned in commenting on the other two tests and the criteria.

Now if two types of test are to be checked for reliability and validity, then it is necessary to choose examples of each test that have been adequately prepared. Daniels does not discuss the question of whether the tests should have equal testing time, or equal construction, testing and marking time. Yet this is obviously important, for with small numbers of candidates, a saving in marking time does not compensate for longer preparation time, but if the number of papers to be marked is large, the saving in marking time can be devoted to test construction.

The first step in preparing an adequate objective test is to draw up a test specification based on the objectives to be tested. Only after his teachers had clearly formulated the aims of geography teaching at Form IV level and specified the kinds of knowledge and abilities they expected their pupils to display should Daniels have begun to construct his objective test. Then, after the items had been prepared according to specifications, they should have been given to a sample of students and an item analysis made to eliminate unsatisfactory items.

Teachers have had plenty of experience in writing essay questions but little in preparing good objective questions. The article does not give sufficient data to judge just how far short of desirable standards the objective test fell. We are not given a mean and standard deviation to indicate the difficulty and ability of the test to discriminate. If the test is easy but speeded a split half reliability coefficient is spuriously high. The total time for a test of 55 factual items would probably be 30-40 minutes, which is almost certainly less than that required to answer three essay questions at this level. The examples given to illustrate the types of question used in the objective test seem to aim at testing only factual knowledge. And why should one-quarter of the test consist of true-false questions, which are usually regarded as among the less satisfactory types of objective questions? An experiment comparing a third-rate example of an 'objective' test against a reasonable example of an essay examination can hardly be expected to supply a valid answer to the question asked by the author.

In establishing the criterion of geographical ability, the teachers appear to have received little detailed guidance. The fact that teachers weighted different aspects of geography differently is clear from Table V, which reveals the weight for the different tests for predicting within schools. The teacher at School X thinks that what is tested by Test A is important; the teacher at School Z apparently thinks the information tested by Test B is of major importance. When given a free hand, teachers would arrive at an estimate which correlates highly with their own examining procedures and in most cases these would be essay tests. At the very least, some analytic analysis should have been made before establishing a single criterion score. The fact that the validity coefficient approaches the reliability coefficient of the essay test suggests that both measure very much the same thing, namely, the ability of the student to do geography questions of an essay type. The degree of care taken in marking the essay questions could be important. Would the markers take more care in marking 30 papers for a research study than in marking hundreds of

papers in a G.C.E. examination? (The writer here assumes that external examination procedure in England does not differ markedly from that in Australia.)

If the University of Nottingham Institute of Education wishes to conduct a useful study of examinations in geography, the following procedure is suggested:

- (1) Get the teachers of East Midlands to discuss their aims in teaching geography at Form IV level.
- (2) When these aims have been clarified take each aim in turn and discuss how progress in attaining that aim can best be measured. In some cases the best type of test will be obvious, e.g., if certain facts are expected to be known then objective tests may be obviously more suitable.
- (3) Where the best method of testing is in any way doubtful, then an effort should be made to construct objective tests as well as essay tests. The task of assessing the relative merits of each in this situation is difficult. For a small number of candidates, the time required for the preparation of the tests by the teachers may be an important factor, but as the number of candidates increases, it becomes possible to spend more time on test construction if time is thus saved on marking. However, it would appear useful to give to a group both essay and objective tests, using the same testing time and prepared so as to measure the same objective. The criterion of geographical ability could be teachers' ratings, provided that the raters took no part in setting or marking either paper, but did take part in formulating the objective.
- (4) If some goals cannot be tested by objective or semi-objective tests, but can be tested by essay tests, then essay tests should continue to be used.

REPLY BY J. C. DANIELS.

S. S. Dunn's interesting comment on my article illustrates the difficulties that are likely to arise when, in the interests of brevity, complete details of an investigation, including full copies of the tests used, are not given.

I am in complete agreement with his suggested procedure for making a 'useful study of examinations in geography.' But the truth is that, except for one minor point, he describes exactly the procedure actually adopted in our investigation. The investigation arose out of lengthy and detailed discussions with the teachers of geography in grammar schools in the East Midlands on the aims of teaching geography at fourth and fifth form levels. One of the points arising out of these discussions was the decision to carry out an empirical investigation to discover the best ways of measuring how far these aims were being achieved.

The three types of test, including the maps and diagrams test, were suggested by the teachers as tests likely to be most useful in measuring how far their aims were being achieved. These origins of the investigation also account for the fact that equal time was not given to the three tests, for the teachers, albeit indirectly, decided what would be a "fair" allocation of time for the different tests. Both the maps and diagrams test and the objective test were tried out on a group of similar pupils not taking part in the experiment and the final versions were drawn up by selecting the best items after a careful item analysis.

The suggestion that the split-half reliability may be spuriously high because the 'objective' test may have a low standard deviation is not supported by the facts. The time taken in the 'objective' test (30 minutes) was ample time for all the pupils to finish the test. For these reasons, the suggestion that our 'objective' test may have been a third-rate example of such tests is, therefore, respectfully denied.

Dunn does, however, seem rather confused about the criteria to be used in estimating the validity of tests—he has to agree to my using teachers' estimates but does so most grudgingly as if hankering after computing the validity of tests against more tests. But he does finally agree with the procedure we adopted.

I ought to make it quite clear that in my subjects and in all situations. Far from discourage the use of 'objective' tests in all subjects and in all situations. Far from it. But it does seem important that some of the rather rash claims of the whole-hog supporters of 'objective' testing as an examinational panacea should have been refuted by this investigation.

SUMMARIES OF RESEARCHES REPORTED IN DEGREE THESES¹

Provisional Admission to Senior Secondary Courses

By JAMES WALKER.

*(Abstract of a thesis submitted in part fulfilment of requirements for the degree of Ed.B.
in the University of Aberdeen, 1955.)*

AIM.

By means of a follow-up, to examine the value of a scheme which provisionally admits to a senior secondary school (the approximate Scottish equivalent of the secondary grammar school) certain pupils who obtain border-line scores in the transfer examination.

SUBJECTS.

- 154 pupils admitted to one Scottish senior secondary school in 1951, including :
- (a) 107 pupils whose scores in the transfer examination were above the border-line.
 - (b) A border-line group of twenty-four pupils, admitted under the established transfer scheme as a result of the further assessments of all border-line pupils ('border-line admits').
 - (c) Twenty-three pupils whose scores in the transfer examination were in the border-line group but whose provisional admission to senior secondary courses was the result of individual consideration and reports after appeal by parents or headmasters ('provisional admits').

CRITERION.

Several criteria of success in the senior secondary courses were adopted : (i) entry to fourth year of study in the school (after age 15) ; (ii) teachers' assessments of suitability for admission to fourth year work (the stage at which specialized preparation for the Scottish Leaving Certificate begins) ; (iii) allocation to higher classes within the school ; (iv) marks in class examinations.

RESULTS.

(i) The proportions of each group entering fourth year of study were : of the provisional admits, 61 per cent. ; of the border-line admits, 21 per cent. ; of the rest, 53 per cent.

(ii) The proportions of each group considered by teachers at the end of the third year to be suitable for fourth-year work were : of the provisional admits, 65 per cent. ; of the border-line admits, 38 per cent. ; of the rest, 50 per cent.

(iii) At the beginning of the third year, eleven of the twenty-three provisional admits had been promoted to a higher class (or stream) ; only three of the twenty-four border-line admits had been similarly promoted.

(iv) Though various classes sat different examinations and thus marks could not be equated, a rough estimate of performance was obtained from a crude average mark in all first-year and third-year examinations. The average percentage marks in all first-year examinations were : provisional admits, 65.4 ; border-line admits, 61.8 ; rest, 67.5. The averages in the third year were : provisional admits, 60.2 ; border-line admits, 56.6 ; rest, 60.5.

(v) A check on attendance during the first two years showed that the provisional admits had fewer absences on average than the other groups.

CONCLUSIONS.

The subsequent performance of those provisionally admitted seems to justify their admission. The result may be due to parental interest, shown in the initial appeal and in pupils' subsequent good attendance, or to the provisional basis of admission, or to the individual consideration given in allocation. Too much must not be read into the result, however, because of the smallness of numbers and because final conclusions must await completion of full five or six-year courses.

¹ Theses Outlines must be submitted through the Head of the Department in which the research was carried out.

The Prognostic Value of the University Bursary Competition

By ELSPET M. YOUNIE.

(Abstract of a thesis submitted in part fulfilment of requirements for the degree of Ed.B. in the University of Aberdeen, 1955.)

Aim.

To assess the validity of the Aberdeen University Bursary Competition (the Entrance Scholarship examination) by means of a follow-up of the subsequent careers of bursary winners and others among the 1933 and 1934 entrants to the Arts Faculty.

SUBJECTS.

The sixty-eight winners of places on the 1933 and 1934 Arts bursary lists (31 men, 37 women) and the 175 others who did not win places (88 men, 87 women), all being entrants to the Arts Faculty immediately subsequent to the Competition.

METHOD.

The academic careers of both groups were traced from University records. Careers after graduation were traced from General Council records and from answers to questionnaires returned by 50 per cent. of bursary winners and 49 per cent. of the others.

RESULTS.

(i) Bursary winners achieved greater academic distinction than the others (chi-square significant at .001). 35 of the 68 bursars graduated with honours (15 first-class); 34 of the 175 others graduated with honours (4 first-class).

(ii) Participation in extra-curricular activities (of those who returned the questionnaires) was greater among bursary winners than among the others; but the bursary winners were most active in societies related to studies (this being due to the higher proportion of honours students among bursary winners), and the others were most active in sports.

(iii) Among honours and ordinary graduates separately, bursary winners were younger on average at graduation than the others (honours, not significant; ordinary, significant).

(iv) In a classification of fathers' occupations, there was no significant difference between bursary winners and others.

(v) Of those who returned questionnaires, bursars on average achieved higher success in their professional careers than the others, but this seemed principally due to superior advancement in teaching, associated with the higher proportion of honours graduates among the bursars. Follow-up of other aspects (honours, publications, etc.) produced only rare mentions from all who returned questionnaires.

(vi) A slightly higher proportion of bursars than the others were married, bursars tending to marry at a younger age than the others (difference not significant). There was no difference in the size of families of those married in the two groups.

CONCLUSION.

The results generally support the validity of the Bursary Competition as a prognostic measure of academic success.

BOOK REVIEWS

Growing Up in the City: A Study of Juvenile Delinquency in an Urban Neighbourhood: J. B. MAYS. (Liverpool: University Press of Liverpool, 1954, pp. x+216, 17s. 6d.)

Mr. Barron Mays, Warden of the Liverpool University Settlement, has selected for intensive study eighty lads who were members of a youth club in one of the poorest quarters of the city. Each was observed and regularly interviewed during a period of a year or more. Thirty had already been convicted of larceny or the like; another twenty-two eventually confessed to similar offences. Mr. Mays' main conclusion is that all the lads were just 'average youths,' and that their lawless behaviour was really "a symptom, not so much of maladjustment as of adjustment to a particular sub-culture"—a culture in which "delinquency has become almost a social tradition." But why do some individuals succumb, whereas others do not? Mr. Mays apparently distrusts any explanation in terms of inborn differences in temperament or intelligence; and is inclined to accept Dr. Bowlby's hypothesis—namely, that those who, during their infant life, have been deprived of maternal affection are themselves inclined later on to develop into "affectionless personalities"—a hypothesis several recent investigators have vigorously questioned.

Perhaps the most instructive parts of his book are those in which he compares the results of his own inquiry with the conclusions of previous investigators. Liverpool, more than any other British city, has furnished opportunities for investigating juvenile delinquency along scientific lines. As long ago as 1908, with the willing co-operation of the police and the education office, I myself attempted to apply what were then regarded as "the newfangled methods of experimental psychology" to the study of youthful offenders referred for examination by the schools or police courts; and, with the help of Frederick Marquis (now Lord Woolton), who was at that time Warden, I was able to make the University Settlement, where I resided, my main headquarters, and use the sociological information acquired by its various members to throw light upon my problems. After the first world war, Mr. J. Bagot, another member of the Social Science Department of the University, undertook a statistical study of court records, which was fully reviewed in this *Journal* (XI, pp. 138-142). Whereas my own investigations were concerned mainly with children of school age, Mr. Mays has been interested chiefly in youths aged 16 to 19; and Dr. Bagot has attempted to cover both.

Mr. Mays considers that the clinical or statistical approach adopted by his predecessors missed some of the most essential factors. "Bagot's inquiry," he says, "is disappointing and inadequate in that it lacks psychological depth and insight into the dynamic character of much delinquent behaviour: . . . it is probably inevitable that an approach based largely on statistics should lack the human element." My own investigation he criticizes for the opposite reason: my "bias (he says) was psychological," and hence I tended to "place psychological causes before social causes": as a result, I "underestimated the influence of the external environment and of bad social traditions in a slum area." In point of fact, I devoted a whole chapter to "environmental conditions outside the home," and concluded that "the character of the street and neighbourhood in which the child lives may be quite as significant as those inside his home or lodging." Mr. Mays, however, believes that I also underrated the influence of the home itself. Quoting Bowlby, he declares that "Burt had previously stumbled on the factor of parental separation, but without realizing its full significance," since (with one or two exceptions) I "placed these early separations among the *minor* factors" (Mr. Mays' italics): such an assignment, he contends, is contradicted by the frequency with which, according to my table, these conditions were noted among the delinquent boys and girls. But here there is a plain misunderstanding. The decision whether a factor played a minor or a major part was based, not on its frequency, but on the case-history. In most instances the parents' treatment of the child in earlier years was merely one of several predisposing factors, not a major or precipitating cause. But in any case, my final conclusion was that, on the whole, "defective home discipline is the most

important of all causal factors," and that, among the older delinquents, the influence of companions or associates "ranks as one of the biggest."

After all, therefore, except for his doubts about the influence of innate tendencies, there is really little difference between us. Mr. Mays explicitly accepts my chief contentions, namely, that "causation is nearly always multiple," that "the family is the place where maladjustment and delinquent behaviour usually begin," and that, save for a few exceptional cases, the delinquent is in no way "abnormal in a psychiatric sense." The minor discrepancies that he has noted are sufficiently explained by the disparity in age between the groups with which we were concerned. And the special merit of his investigation is that it shows how much the same conditions are conducive to criminal conduct among older lads as among boys and girls of school age. The whole inquiry forms a valuable contribution to the subject, based on first-hand practical experience and a sympathetic insight into the problems of contemporary youth, and deserves the fullest attention from teachers, social workers, public authorities, and all who are interested in the welfare of the younger generation.

CYRIL BURT.

English Social Differences: T. H. PEAR. (London: Allen and Unwin, 1955, pp. 318, 18s.)

In this book Professor Pear sets out to do a piece of 'natural history' on England. He collects a number of specimens to illustrate English social differences; and he discusses some of the concepts which may be useful for this purpose. Thus, he picks out the differences exhibited in ways of speaking, manners, etiquette and goodwill, in clothes and fashions, in eating and drinking, in snobbery and in leisure pursuits. The concepts that he discusses in particular are class, stratum, status, and élite. He decides that the concept of class is "insufficient to subserve the chief facts," and that the others "are useful for different purposes." The means he uses to collect his specimens are various. He reports and brings together the work of other investigators. He reports some of his own specimen hunting. He notes the expressed attitudes and opinions of various people. He records many of his own observations and hunches—in manner of a bird-watching field worker. He hopes that, if two types of psychology have been distinguished, that of "understanding" (*verstehende*) and the "scientific" (*wissenschaftliche*), this work will make a modest contribution to the psychology of understanding.

No doubt the two chapters that will be of most interest to readers of this journal are those in which Pear discusses "The Training of Character in some English Schools," and "Social Differences in English Education." In the former Pear considers the traditional defence of the residential "public school" for boys, namely that they are especially good for the training of character and leadership. Pear says something about the concept of "character," notes (in the next chapter) that we lack a social psychology of masters and pupils written by someone who "empathises" their special culture pattern, is mildly sceptical of the results claimed for the public schools, and quotes relevant claims and counter-claims about these schools by various writers. I confess that I do not think Pear has gone the right way about this topic. When a man who has been through a public school claims that these schools are particularly good at training character, Pear treats this as an empirical claim—as a statement of psychological fact. But I suspect that the public-school apologist has no wish here to talk psychology at all (note the resistance when you suggest making a scientific enquiry into the truth of his claim!); and that the chief point of this statement is self-congratulatory. The apologist is expressing, in a conventionally permissible way, a laudatory view of himself—"Look what a fine chap I am! I'm the man to tell others what to do." If I am right, then it is misleading to take this claim for the public schools seriously, as Pear does. It is perhaps more worth while to note that this sort of claim is just another example of social differences. Naturally the public school man makes it for *his* schools, since this is consistent with his role as a member of the dominant groups in our community. Naturally enough the day school man will tend *not* to make this claim for his schools,

since this would be inconsistent with his role as a member of the non-dominant groups in our community. I am inclined to think that it might have been more profitable if Pear had worked out the sort, or sorts, of person likely to be turned out from the paternalistic-cum-authoritarian atmosphere of a public school, with its prefects, fagging, compulsory chapels, dormitory prayers before house matches, caning, etc. If he had then contrasted this sort/s of person with the sort, or sorts, likely to be produced from differently organised schools, he could then have gone to the facts to discover whether they bear out his expectations or not. If the known facts are insufficient for this purpose, Pear will then have shown us the sort of thing we need to know in order to decide whether differences in schooling produce differences in personality and character.

In his chapter on "Social Differences in English Education," Pear discusses, in particular, the social stratification of the English University system by reference to the role of Oxbridge and its contrast with Redbrick. He argues that Oxbridge has in the past served to embody this social stratification, and he seems to imply that it still does so, in spite of the great increases in educational opportunity since the war. Yet he throws doubt on the "cultural advantages" that Oxbridge is alleged to bestow by quoting from a critical letter by a man who went up to Oxbridge in 1925. He attempts a comparative picture of Oxbridge and Redbrick by quoting a report of a woman who was at Oxbridge from 1939-42, and at Redbrick in 1942-43; and he adds a comment on this report by a woman graduate who was at Cambridge in 1949. I am forced to say that I am not happy about all this. It is not much use quoting a report by a man who came up to Oxbridge in 1925 if one is interested in the advantages that Oxbridge can bestow *at present*. The drive by schools to get pupils into Oxbridge, rather than Redbrick, may not in itself be evidence of *social* stratification so much as evidence of a belief among school teachers in *academic* stratification, namely that Oxbridge provides a better education than Redbrick. If Pear wishes to throw doubt on this belief, the obvious way to do so is to question the value of the tutorial system, a feature of Oxbridge that he does not appear to consider. Moreover, the post-war changes have been vast. In particular, I am given to understand that the great majority of undergraduates now come up on public funds. Pear does not consider what social effects, if any, these changes have produced. It will not do to rely on two impressionistic reports from the years 1939-43 and 1949.

This leads me on to confess that I am uneasy about the book as a whole. If one is to do a piece of natural history in the psycho-social field, it is advisable to do so systematically, and to be quite clear about the methods one is using, their value and limits. I do not think that Pear passes on these two counts in this book. Moreover, if one asks: "What new information does this book contain?" the answer is: It does not tell us much that the ordinary educated person did not know already. If one asks: "Does it make a theoretical contribution by, e.g., its analysis of concepts?" the answer is again: No, the analysis presented of concepts is too brief and sketchy for this purpose. And all this is unfortunate if Pear wishes to impress the influential groups in our community with the important contribution which psychology has to offer.

But does the book make a contribution to the psychology of *Verstehen*, as Pear hopes? I think it does, in so far as it alerts us to notice social differences that we would otherwise have missed. But I doubt whether it contributes to the understanding of the English in the rich and intimate way in which the work of a social anthropologist like, say, Richards, enables one to understand the Bemba of Northern Rhodesia during the 1930's (see, e.g., *Bemba Marriage and present Economic Conditions*, by Audrey I. Richards: Rhodes-Livingstone Institute, papers No. 4, 1940). On the other hand, what Pear's book also does, I think, is to help break the taboo on the discussion of some of these issues of social differences; to show us that these are more complex than we were perhaps inclined to suppose; and to emphasise just how little we know on these interesting and important matters. I hope his work will stimulate others to carry on where he leaves off.

B. A. FARRELL.

Deprived Children: A Social and Clinical Study: HILDA LEWIS. (Oxford University Press, pp. 163, 9s. 6d.)

It is a great relief to deal with a book concerned with deprived, deserted, or otherwise problem children, in which neither excessive sentiment nor the over emphasis of one or two supposed causes of the trouble spoil the argument, or make a critical reader feel that he cannot trust the judgment of the writer. Dr. Hilda Lewis, while revealing a deep concern for the welfare of these unfortunate children, is careful to give detailed statistics on many aspects of the home conditions of the children and of their own personalities, and she is rightly cautious in drawing conclusions.

The book deals with 500 children who were admitted to the Reception Centre at Mersham, Kent, for observation for a short period, before it was decided whether they should be placed in foster homes or institutions, or returned to their families and so forth. On the whole these children would be the worst examples of the many unfortunate children for whom the Kent County Council became responsible. About half the children remained at the centre between three and six weeks; most of the rest were there less than three weeks.

A follow-up study was made in most cases about two years after placement, partly by the visits of an investigator, and partly by written reports. Numerous tables (there are seventy-one) give precise details as to the condition of the children at reception, the nature of the home and the reports two years later. It is gratifying to find a psychiatrist who does not shirk precise statistical estimates, including even the use of χ^2 . While Dr. Lewis confirms the view as to the tendency for children to suffer from the deprivation of mother love, she stresses the fact that early separation from the mother need not necessarily have serious effects, as has been so confidently stated. Thus: "Separation from the mother before the age of 5 years was a prognostically adverse feature. Nearly a third of the children who were separated from their mothers were, however, in a satisfactory condition at the end of the follow-up period, and permanent separation before the age of 2 had not been the prelude to a particularly unsatisfactory condition of the child at the end of the period."

Further quotations illustrate the author's broad outlook, including in particular the admission of the importance of hereditary factors so often ignored. It was found that 82 per cent. of the mothers and 81 per cent. of the fathers were suffering from mental disabilities—neurosis, psychosis, psychopathic personality or mental defect or dullness, and Dr. Lewis comments: "It would be, indeed, strange if so many maladjusted, dull, neurotic, and psychopathic parents as are included in this series did not hand on some of their defects to their children." Again she writes: "More than a fifth of all the children admitted to a reception centre chiefly because they had no satisfactory home and background, nevertheless were normal in their behaviour and gave no grounds for inferring that psychological damage had been done."

This is certainly a book which no serious student of child welfare or of maladjusted children should neglect. C.W.V.

The Application of Psychological Tests in Schools: N. E. WHILDE. (London: Blackie, 1955, pp. 179, 9s.)

This is an elementary book, designed for teachers. It covers intelligence and attainment testing in some detail, and looks briefly at performance tests, personality tests, rating scales, etc. Many of the staff-room criticisms of tests are considered and met, and the advice on the use of tests is, on the whole, sound and helpful. The author's experience and interests (in child guidance) are to some extent reflected in the relative emphasis laid on different topics. For example, 'reliability' receives a most cursory treatment, although a fuller development of this, leading to a discussion of standard error of score, would have illuminated the section on 'constancy of the I.Q.' and helped to prevent the not uncommon belief in the accuracy of quotients to within one point. The author quite rightly says, in his preface: "Tests suffer from those who claim too much for them as well as from those who are prejudiced

against them." It is probable that, for books at this level, it is more important to prevent an increase in numbers in the former category than to attempt to reform those already in the latter. For this reason alone, it is unfortunate that the list of test material available contains such things as the Lowenfeld Mosaic and Kaleidobloc tests and Burt's Questionnaire on Neurotic Symptoms. In any case, a long list of intelligence and attainment tests such as this does little to help the teacher to choose the test which really suits his purpose. Even the time-lengths are given for only a few of them. A shorter list, with fuller informative (and critical) notes, would be more useful.

Probably this book would be most effectively used on teachers' courses, to supplement other reading such as Vernon's *Measurement of Abilities*, Dempster's *Selection for Secondary Education*, and Richardson's *Mental Measurement*.

STEPHEN WISEMAN.

An Introduction to Mental Measurement and its Applications: C. A. RICHARDSON. (Longmans, 1954, pp. 102, 8s. 6d.)

This is an attempt to explain the principles of mental testing and their applications in school to teachers and students who have no special aptitude or interest in mathematics. The book covers familiar ground, being divided roughly equally between the statistics of mental measurement and the application of tests in education and in vocational guidance. The first half deals with such concepts as standard deviation, percentile, significance, correlation coefficient and regression, using only the simplest arithmetic to illustrate them. It is very clearly written and a student who is undergoing a course in elementary statistics will find much here which will clarify his thought. It is doubtful, however, whether many readers with no statistical background will be able to understand these (to him) novel concepts without the experience of them in numerous contexts; this will be specially true of the chapters dealing with correlation and factors of the mind.

The second half of the book is a clear, modest statement of the uses and limitations of tests in schools for classification purposes, for comparing standards and for detecting the underfunctioning child. There is a sound chapter on allocation to secondary schools and a cautious survey of the present position in the use of tests for vocational guidance.

The book will probably find its main use in the hands of teachers and others who, having some background knowledge, wish to bring their knowledge up to date.

L. B. B.

Experimental Psychology: Editor, B. A. FARRELL. (Blackwell, 1955, p. xi+66, 7s. 6d.)

This is a group of talks broadcast in the Third Programme last year and aimed at giving an informed non-psychological audience a glimpse of the field in which experimental psychologists are working and an idea of the methods they employ and the attitudes they develop in studying psychological phenomena. This is done by taking a sample of workers in different aspects of experimental psychology, mainly from Oxford, and allowing each to talk upon a topic of special interest to him. Each talk is thus a unit in itself, and the topics, which include Watson on Perception, Kay on Learning, Deutsch on Motivation, Farrell on Psycho-Analysis and Argyle on Social Behaviour, are well chosen to show the great range of subjects which are being tackled. Though there is little here which is specially relevant to educational psychologists, there is no doubt that they will enjoy reading these talks and some may benefit from them by being provoked to seek new ways of dealing with their own problems.

L. B. BIRCH.

The Forge, The History of Goldsmiths' College: Edited by DOROTHY DYMOND. (Methuen, pp. 158, 21s.)

This book will appeal not only to old students of Goldsmiths' College, but to many others interested in the history of the training of teachers. The two main contributors are well-known personally in connection with the education of teachers, viz.:—Principal, A. E. Dean and Miss Caroline Graveson.

Devon Interest Test: S. WISEMAN and T. F. FITZPATRICK. (Oliver and Boyd, pp. 21, Test 8d., Manual of Instructions, 2s.)

These tests, of which there are separate ones for boys and girls, should provide useful material for experiments in the selection of pupils for grammar schools *versus* technical schools. The *Manual of Instructions* is more than that; it also gives details of several applications of the test, though the authors frankly state that the test has not yet been "fully validated by follow up." C.W.V.

OTHER PUBLICATIONS RECEIVED

The mention of a book in this list does not preclude a later review.

- Kant*: S. KORNER. (Pelican, pp. 230, 2s. 6d.)
- The Young Worker*: Report of Fifth Conference at Oxford, April, 1954, University Department of Education. (Heinemann, pp. 64, 3s. 6d.)
- Promoting Maximal Reading Growth Among Able Learners*: H. M. ROBINSON. (Univ. of Chicago Press, 26s., pp. 191.)
- American Indian and White Children—A Socio-psychological Investigation*: R. J. HAVIGHURST and B. L. NEUGARTEN. (Univ. of Chicago Press, pp. 335, 37s. 6d.)
- L'Utilisation des Aptitudes*: H. PIERON and OTHERS. (Presses Universitaires de France, pp. 757, 1,500 frcs.)
- The Teaching of Nature Study*: T. Y. HARRIS. (Australian Council for Educational Research, Melbourne, pp. 100, no price given.)
- Principes des L'éducation Religieuse, Morale et Sociale*: A. KRIEKEMANS. (Editions Nauwelaerto, Louvain, pp. 164, 69 frcs.b.)
- The Child's Construction of Reality*: JEAN PIAGET. (Routledge and Kegan Paul, translated from the French by Margaret Cook, pp. 389, 25s.)
- Hunter College—Eighty-five Years of Service*: S. W. PATTERSON. (Lantern Press, New York, pp. 263, \$3.50.)
- Beginning to Read*: E. R. BOYCE. (Macmillan and Co., Ltd., pp. 32, 3s.)
- Notes on Geography*: A. W. HARDING, B.A. (Blackwell, pp. 20, 1s.)
- Notes on Punctuation*: ERIC PARTRIDGE. (Blackwell, pp. 20, 1s.)
- Notes on Chemistry*: C. K. HILLARD, M.A. (Blackwell, pp. 24, 1s.)
- Notes on History*: C. R. N. ROUTH, M.A. (Blackwell, 1763-1846, Book I, pp. 20, 1s.)
- Notes on History*: C. R. N. ROUTH, M.A. (Blackwell, 1846-1914, Book II, pp. 20, 1s.)
- Liberal Education in a Technical Age*: NATIONAL INSTITUTE OF ADULT EDUCATION. (Max Parrish, London, pp. 128, 6s.)
- An Introduction to Psychology*: HAIRY W. KARN and JOSEPH WEITZ. (Chapman and Hall, pp. 315, 32s.)
- The Learner on Education for the Professions*: CHARLOTTE TOWLE. (Chicago Univ. Press, pp. 432, 56s. 6d.)
- La Psicologia della Eta Evolutiva—Quarta Edizione*: FR. A. GEMELLI. (Giuffrè, Milano, pp. 379, lira 2,000.)
- Australian Journal of Psychology—Monograph Supplement No. 1, Melbourne Test 90*: PAUL LAFITTE. (Melbourne Univ. Press, pp. 107, 12s. 6d.)
- Low Intelligence and Delinquency*: MARY WOODWARD. (Institute for the Study of Delinquency, pp. 24, 1s. 6d.)
- Educational Measurement*: ROBERT M. W. TRAVERS. (Macmillan Co., New York, pp. 420, \$4.75.)
- An Inquiry About Examination* (Calcutta University): NIROD MUKERJI. (Das Gupta and Co., Calcutta, pp. 67.)
- Psychological Statistics* (Second Edit.): QUINN MCNEMAR. (Chapman and Hall, pp. 408, 48s.)
- I, Problemi Della Pedagogia*: Direttore LUIGI VOLPICELLI. (Insituto di Pedagogia Dell 'Universita' di Roma, L.800.)

- Hawera—A Social Survey*: Editor A. A. CONGALTON. (Hawera and District Progressive Assn., pp. 218, 10s.)
- Análise das Profissões Comerciais—O. Balconista—Monograph 2*: L. A. COSTA PINTO. (Departamento Nacional do Senac, pp. 62, no price given.)
- Die Regulierungs-Kräfte-Im Seelen-Leben*: FRANZISKA BAUMGARTEN. (A. Francke Ag. Verlag, pp. 138, Fr. 2-80.)
- To Define True Madness*: HENRY YELLOWLEES. (A Pelican Book. Penguin Books, Ltd., pp. 172, 2s. 6d.)
- Thorndike School Dictionary*: E. L. THORNDIKE. (Univ. of London Press, pp. 371, 5s.)
- De La Logique de L'Enfant a la Logique de L'Adolescent*: B. INHELDER and JEAN PIAGET. (Presses Universitaires de France, pp. 314, Frs. 1,000.)
- A Brief Guide to Australian Universities*: Third Edition. Australian Council for Educational Research. (Melbourne Univ. Press, pp. 42, 4s.)
- The Academic Record of Science Students in the University of New Zealand*: Bulletin 113: I. D. DICK, R. M. WILLIAMS, and D. STRAKER. (New Zealand Depart. of Scientific and Industrial Research, pp. 64, 6s. 6d.)
- The Golden Monkey (Simplified Edition)*: CAPT. F. KNIGHT. (Macmillan, pp. 140, 5s.)
- Management of Addictions*: E. PODOLSKY, M.D. (Philosophical Library, New York, pp. 413, \$7-50.)
- Delinquent Boys, the Culture of the Gang*: A. K. COHEN. (The Free Press, Glencoe, Illinois, pp. 202, \$3-50.)
- Truants from Life*: B. BETTELHEIM. (The Free Press, Glencoe, Illinois, pp. 511, \$6-00.)
- La Formation du Concept de Réflexe: Aux XVIIe and XVIIIe Siècles*: G. CANGUILHEM. (Presses Universitaires de France, pp. 206, Frs. 600.)
- Education and Psychology, Mental Testing Number, 1954, 4-5*. Editors, PRAYAG MEHTA, and UDIA PAREEK. (J. M. OJHA, Calcutta, pp. 180, Sub. Annual Rs. 12s.)
- The Scottish Pupil's Spelling Book, Parts I to V*: Prepared by Scottish Council for Research in Education. (Univ. of London Press, pp. 29-48, 1s. each.)
- The Psychology of the Criminal Act and Punishment*: GREGORY ZILBOORG, M.D. (Hogarth Press and Institute of Psycho-Analysis, pp. 150, 10s. 6d.)
- Clinical Papers and Essays on Psycho-Analysis*: KARL ABRAHAM, M.D. (Hogarth Press and Institute of Psycho-Analysis, pp. 336, 25s.)
- Hospitalization of Mental Patients: A Survey of Existing Legislation*. (World Health Organization, Palais des Nations, Geneva, pp. 100, 6s. 9d.)
- Alcohol and Alcoholism, Report of an Expert Committee*. (World Health Organization, Geneva, pp. 14, 1s. 9d.)
- Adventure in the River*: C. V. BURGESS. (Univ. of London Press, pp. 32, 1s. 9d., Work Book, 1s.)
- The Soap-Box Derby*: C. V. BURGESS. (Univ. of London Press, pp. 32, 1s. 9d., Work Book, 1s.)
- Adventure on the River*: C. V. BURGESS. (Univ. of London Press, pp. 32, 1s. 9d., Work Book, 1s.)
- George Takes the Lead*: C. V. BURGESS. (Univ. of London Press, pp. 32, 1s. 9d., Work Book, 1s.)
- The Adolescent Years*: WILLIAM W. WATTENBERG. (Harcourt Brace and Co., New York, pp. 510, \$3-55.)

INDEX TO VOLUME XXV, 1955

Articles :	PAGE
Attainment in Primary Schools, Environmental and other Characteristics Determining : LESLIE D. C. KEMP	67
Attainments in England and Scotland, A Comparative Study of Education : P. E. VERNON, M. B. O'GORMAN and A. McLELLAN	195
Co-Education : The Verdict of Experience : R. R. DALE	145
Composition, English, in Secondary School Selection : J. D. NISBET	51
Concept of Intelligence, The Evidence for the : SIR CYRIL BURT	158
Feeble-Mindedness, An Investigation Bearing on the Laws Relating to, and their Application : F. K. FITZPATRICK	117
Group Research with Post-Graduate Student Labour in a Department of Education : E. W. HUGHES	99
Individual Activities in the Junior School, Children's Choices in : E. W. HUGHES	36
Intelligence and Attainment Tests as Predictors of Success in Grammar Schools, the Relative Efficiency of : JACK WRIGLEY	107
Interest Test in 11 plus Selection, The Use of : STEPHEN WISEMAN	92
McDougall, The Permanent Contribution to Psychology of : SIR CYRIL BURT	10
Mathematics, A Study of Specific Ability and Attainment in : D. M. LEE	178
Nail-Biting Among School Children, the Incidence of : L. B. BIRCH	123
Newspaper Reading in the Secondary Modern School : D. G. BURNS	1
Premature Leaving from Grammar Schools, the Causes of, Part II : M. COLLINS	23
Science, Psychology and the Teaching of : E. A. PEEL	135
Teachers in Scotland under the Post-War Emergency Scheme, A Survey of Students admitted to train as : JOHN SUTHERLAND	78
Terman-Merrill and Matrices Test Data, the Validity and Interchangeability of : D. WALTON	190
Testing Geography, A Criticism of the Article by J. C. Daniels : S. S. DUNN	204
Testing Geography, A Reply to Mr. S. S. Dunn's Criticism : J. C. DANIELS	205
Thomson, Sir Godfrey, An Obituary Notice : JOHN SUTHERLAND	65
Undergraduates, Ex-National Service and " Schoolboy "—A Comparative Study of Academic Performance of : E. M. BROWN and J. G. M'COMISKY	55
Summaries of Researches Reported in Degree Theses	129, 206
Authors :	
BIRCH, L. B. : The Incidence of Nail-Biting among School Children	123
BROWN, E. M. and M'COMISKY, J. G. : Ex-National Service and " School-boy " Undergraduates : A Comparative Study of Academic Performance	55
BURNS, D. G. : Newspaper Reading in the Secondary Modern School	1
BURT, SIR CYRIL : The Permanent Contribution of McDougall to Psychology	10
BURT, SIR CYRIL : The Evidence for the Concept of Intelligence	158
COLLINS, M. : The Causes of Premature Leaving from Grammar Schools, Part II	23
DALE, R. R. : Co-education : The Verdict of Experience	145
DANIELS, J. C. : A Reply to Mr. S. S. Dunn's Criticism	205
DUNN, S. S. : A Criticism of the Article by J. C. Daniels on " Testing Geography "	204
FITZPATRICK, F. K. : An Investigation Bearing on the Laws Relating to Feeble- Mindedness, and their Application	117
HUGHES, E. W. : Children's Choices in Individual Activities in the Junior School	36
HUGHES, E. W. : Group Research with Post-Graduate Student Labour in a Depart- ment of Education	99
KEMP, LESLIE, C. D. : Environmental and other Characteristics Determining Attain- ment in Primary Schools	67
LEE, D. M. : A Study of Specific Ability and Attainment in Mathematics	178
NISBET, J. D. : English Composition in Secondary School Selection	51
PEEL, E. A. : Psychology and the Teaching of Science	135
SUTHERLAND, JOHN : A Survey of Students Admitted to Train as Teachers in Scotland under the Post-War Emergency Scheme	78
SUTHERLAND, JOHN : Sir Godfrey Thomson—Obituary Notice	65
VERNON, P. E., O'GORMAN, M. B., and McLELLAN, A. : A Comparative Study of Educational Attainments in England and Scotland	195
WALTON, D. : The Validity and Interchangeability of Terman-Merrill and Matrices Test Data	190
WISEMAN, STEPHEN : The Use of an Interest Test in 11-plus Selection	92
WRIGLEY, JACK : The Relative Efficiency of Intelligence and Attainment Tests as Predictors of Success in Grammar Schools	107

